WAITING TIMES FOR CANCEROUS MUTATIONS IN TWO MATHEMATICAL MODELS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Stephen Savinar Moseley

August 2011

WAITING TIMES FOR CANCEROUS MUTATIONS IN TWO MATHEMATICAL MODELS

Stephen Savinar Moseley, Ph.D.

Cornell University 2011

This dissertation explores the distribution of τ_k, the first time for some cell to accumulate k mutations, under two different multistage models for cancer growth. The first model considers a freely mixing, exponentially growing cell population, while the second model considers a spatially fixed cell population of constant size.

The first model is inspired by previous work of Iwasa, Nowak, and Michor (2006), and Haeno, Iwasa, and Michor (2007). We consider an exponentially growing population of cancerous cells that will evolve resistance to treatment after one mutation, or display a disease phenotype after two or more mutations. We use multi-type branching processes to prove results about τ_k and about the growth of the number of type k cells, and apply our results to re-derive proofs in Iwasa, Nowak, and Michor (2006) and Haeno, Iwasa, and Michor (2007) concerning the likelihood of a type k mutant by the time the tumor reaches size M.

The second model is inspired by Komarova (2006). We consider a multi-type Moran model in which cells inhabit the d-dimensional integer lattice. Starting with all wild-type cells, we prove results about the distribution of τ_2 in dimensions d = 1, 2, and 3, and use results from neutral and biased voter models to consider the effects neutral and advantageous mutations, respectively.

BIOGRAPHICAL SKETCH

Stephen Moseley was born in Los Angeles, California on July 12, 1983, and moved with his family to Santa Cruz, California when he was nine. After a childhood spent at the beach and among the redwoods, he headed east to Williams College in pursuit of a Bachelor's Degree in Mathematics, which he received in 2005. While at Williams, he read the complete works of Leo Tolstoy, rowed crew for a year, and developed an undying love for the Boston Red Sox (two years before they broke the curse). The incredible faculty of the Williams Mathematics Department helped deepen his love of applied mathematics, and upon graduating, he rolled right into the Ph.D. program at Cornell University, far above Cayuga's waters.

After ten consecutive years of study, Stephen has finally slaked his thirst for academic mathematics, and has accepted a position with OC&C Strategy Consultants in Boston, Massachusetts. His next great challenge is finding a way to fit Red Sox season tickets into his budget.

Dedicated to my grandfather, Harold Savinar.

ACKNOWLEDGEMENTS

Above all else, I need to thank my adviser, Rick Durrett, whose invaluable contributions made this work possible. I simply could not have completed this dissertation without Rick's generosity, patience, and excellence as a teacher and mentor.

I would also like to recognize the other members of my committee: Dexter Kozen, Gennady Samorodnitsky, and Laurent Saloff-Coste. Six years is a long time to dedicate to any one pursuit; their courses and guidance continuously renewed my love for probability and computer science.

Finally, my deepest thanks and sincerest gratitude to my family and friends. To my parents, Hal and Julie, my sister, Rachel, and my grandparents, Harold and Gloria: your love and encouragement sustained me over my entire academic journey. To my friends, Matthew Barhight, Ryan Belmont, Adriel Cepeda Derieux, Stefan Ragnarsson, Chris Scheper, Eric Sherwood, and my girlfriend (and best friend), Elizabeth Hodgman: you helped me maintain a sense of humor, provided much needed distraction, and gave me perspective when things got tough.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The mathematical investigation of cancer began in the 1950s, when Nordling (1953), Armitage and Doll (1954, 1957), and Fisher (1959) set out to explain the age-dependent incidence curves of human cancers. For a nice survey see Frank (2007). Armitage and Doll (1954) noticed that log-log plots of cancer incidence data are linear for a large number of cancer types; for example, colorectal cancer incidence has a slope of 5.18 in men and 4.97 in women. The authors used this observation to argue that cancer is a multi-stage process, resulting from the accumulation of multiple genetic alterations in a single cell. The math underlying this hypothesis was very simple: suppose X_i are independent and have an exponential distribution with rates u_i (i.e., the density function is $u_i e^{-u_i t}$ and the mean is $1/u_i$). Noting that the sum $X_1 + \cdots + X_k$ has a density function that is asymptotically

$$u_1 \cdots u_k \, \frac{t^{k-1}}{(k-1)!} \quad \text{as } t \to 0, \tag{1.0.1}$$

the authors inferred that the slope of the age-incidence curve was the number of stages minus 1, making colon cancer a six-stage process.

Later on, Knudson (1971) performed a statistical analysis of retinoblastoma, a childhood eye cancer. His study showed that familial cases of retinoblastoma have an earlier age of onset than the sporadic cases that emerge in families without a history of the disease. Based on age incidence curves in the two groups, he hypothesized that two mutagenic events, or "hits," are necessary to cause cancer in the sporadic case, but in individuals with the inherited form of the disease, a single hit is sufficient since one mutation is already present at birth. This study led to the concept of a tumor suppressor gene, i.e., a gene which con-

tributes to tumorigenesis if inactivated in both alleles. See Knudson (2001) for a survey.

The accumulation of several mutations is important not only for cancer initiation and progression, but also for the emergence of acquired resistance against chemotherapeutics, radiation therapy, or targeted drugs. For this reason, Knudson's research led to an explosion of literature on the waiting time τ_k, the first time at which some cell has acquired k prespecified mutations. Most studies, like the ones cited in the last several paragraphs, fit curves of data on age specific incidence without considering a population genetic model for the underlying cell population. Iwasa et al. (2004, 2005) were the first to study waiting times in this way. In their paper, and for all models discussed in this dissertation, type i individuals are cells with i mutations, and type i individuals mutate to type $i+1$ at rate u_{i+1}, which we assume to be much less than 1. Iwasa et al. (2004, 2005) used a Moran model for a population of a fixed size N, considering a variety of scenarios based on the relative fitnesses of mutants. In the neutral case, i.e., if the mutation does not alter the fitness or growth rate of the cell, they showed:

Theorem 1. *In a population of N cells, τ_2 is approximately exponentially distributed with rate $Nu_1u_2^{1/2}$, provided $1/\sqrt{u_2} \ll N \ll 1/u_1$.*

The condition in Theorem 1 ensures "stochastic tunneling", i.e., under these parameters, the first type 2 is almost surely born before the 1's reach fixation. For an intuitive discussion, see Section 12.4 of Nowak's (2006) book. Durrett and Schmidt (2008) applied these ideas to study regulatory sequence evolution and to expose flaws in Michael Behe's arguments for intelligent design. Durrett, Schmidt, and Schweinsberg (2009), see also Schweinsberg (2008), generalized this result to cover τ_k.

In this dissertation, we explore the distribution of τ_k in two different models, each appropriate for different types of cancers. All results come from two papers, which are reproduced verbatim in Chapters 2 and 3. In Chapter 2 we consider τ_k in a freely mixing, exponentially growing cell population, and in Chapter 3, we consider τ_2 in a multi-type Moran model that takes into account cells' spatial positioning. For the remainder of this chapter, we present an overview of each model along with the main results; details and proofs are left for the following chapters.

1.1 An exponentially growing, freely mixing cell population model

In many cases, such as leukemia and other hematological cancers, the cell population is freely mixing and does not have constant size. For these reasons, Iwasa, Nowak, and Michor (2006) considered the time to develop one mutation in an exponentially growing population and Haeno, Iwasa, and Michor (2007) extended the analysis to waiting for two mutations. Their model is a multi-type branching process in which type i cells give birth at rate a_i and die at rate b_i, with $\lambda_i = a_i - b_i > 0$.

In Chapter 2, we consider their model (with a slight modification, detailed below), extending analysis to all τ_k while restricting our focus to cases in which mutation confers a slight fitness advantage, i.e., $i \to \lambda_i$ is increasing.

In Iwasa, Nowak, and Michor (2006) and Haeno, Iwasa, and Michor (2007), type i cells become type $i+1$ with probability u_{i+1} at birth, which translates into a

mutation rate of $a_i u_{i+1}$, while in each of our models, type-i cells mutate at rate u_{i+1} during their lifetimes. This difference must be kept in mind when comparing results, although in applications the mutation rates are very small compared to birth and death rates, so the reduction of the birth rate of type-i's to $a_i(1 - u_{i+1})$ is insignificant.

Let $Z_i(t)$ be the number of type i cells at time t, and $\Omega_\infty^0 = \{Z_0(t) > 0 \text{ for all } t \geq 0\}$, the set on which a new type 0's lineage doesn't die out. We show that

$$(e^{-\lambda_0 t} Z_0(t) | \Omega_\infty^0) \to V_0 = \text{ exponential}(\lambda_0/a_0)$$

and extend a weaker version of this result to all k in Theorem 5:

For $k \geq 2$, $e^{-\lambda_k t} Z_k(t) \to W_k$ a.s. with

$$EW_k = \prod_{j=1}^{k} \frac{u_j}{\lambda_k - \lambda_{j-1}}$$

To uncover more information about the limiting distribution, we let $Z_0^*(t) = e^{\lambda_0 t} V_0$, a close approximation of $Z_0(t)$ continuous for all t (note $Z_0(t) = 0$ for $t < 0$, and jumps to 1 at $t = 0$), and define $Z_i^*(t)$ as the number of type i's at time t in a system where the number of 0's is given by $Z_0^*(t)$ for all $t \in (-\infty, \infty)$.

For $i \geq 2$, we show that $EZ_i^*(t)$ is an upper bound for $EZ_i(t)$, and use typical parameter values to demonstrate that the two are close in application. We are interested in $Z_i^*(t)$ because we can calculate its limiting distribution exactly (see Theorem 6):

For $k \geq 2$, $e^{-\lambda_k t} Z_k^*(t) \to V_k$ a.s. with

$$Ee^{-\theta V_k} = \left(1 + c_{\theta,k} \mu_k \theta^{\lambda_0/\lambda_k}\right)^{-1}$$

Where the constants are given in Chapter 2. From this Laplace transform, we derive our main result:

$$P(\tau_{k+1} > t_{1/2}^{k+1} + x/\lambda_0) \to \frac{1}{1+e^x}$$

Here, τ_{k+1} is the waiting time for the first type $k+1$ in a system with $Z_0^*(t)$ type 0's, and $t_{1/2}^{k+1}$ is the median value of τ_{k+1}, given in 2.1.11. Figure 2.1 compares this result with simulations of τ_3, and we see that despite the approximations, this result accurately models behavior found in simulation.

Finally, in Section 2.1.5, we show that with slight modification our model can be used replicate results from Iwasa, Nowak, and Michor (2006) and Haeno, Iwasa, and Michor (2007) on the likelihood of a type 1 or 2 emerging by times T_M, the first time there are M type 0 cells, and S_M, the first time the total population has reached size M.

1.2 A spatial model for tumor growth

For simplicity, many multistage cancer models assume homogeneously mixing cell populations; however, this is not realistic for many solid tumors. For this reason, Komarova (2006) considered a spatial model which had been introduced much earlier by Williams and Bjerknes (1972), and is the basis of the model studied in Chapter 3. We first consider a version with an infinite number of cells, since it is easier to describe: each location on the d-dimensional integer lattice $\mathbb{Z}^d$ is inhabited by a single cell, initially all of type 0. We suppose that cells of type 0 and 1 have relative fitness 1 and r. Since we only consider the

waiting time for the first type 2 in this chapter, the relative fitness of type 2's is not important. As before, type i cells mutate at rate u_i. At times of a Poisson process with rate 1, each cell x selects one of its $2d$ nearest neighbors y, chosen uniformly at random. If the two cells are the same type, nothing happens, but if they differ, x adopts the type of y with probability $1/(r+1)$ if y is type 0, and $r/(1+r)$ if y is type 1.

When there are no mutations and $r = 1$, this is the voter model introduced independently by Clifford and Sudbury and Holley and Liggett (1975). For a summary of what is known see Liggett (1999). Since we want a finite cell population, we restrict our process to a subset of $[-L/2, L/2)^d$ and denote it by $\bar{\xi}_t^0$. Komarova (2006) uses "Dirichlet boundary conditions", i.e., she assumes her space is an interval with no cells outside, but this is awkward because the set of type 1 cells may reach one end of the interval and then no further changes happen at that end. To avoid this, we will use periodic boundary conditions, i.e., we consider $(\mathbb{Z} \bmod L)^d$. When $d = 2$ this means we are thinking of your skin as a torus, which is a little odd. However, using $(\mathbb{Z} \bmod L)^d$ has the advantage that the space looks the same seen from any point, and our results will show that for the parameter values we consider, the first type 2 will arise when the radius of $\bar{\xi}_t^0$ is $\ll L$ so the boundary conditions do not matter.

In Chapter 3, we are concerned with finding the distribution of τ_2 in the cases of neutral and advantageous mutations ($r = 1$ and $r > 1$, respectively), and in all dimensions $d \geq 1$. To do so, we require the following conditions on parameters.

In a fixed population with $r \geq 1$, the time for type 1's to reach fixation is well understood, so finding the distribution of τ_2 is simple if type 1 fixation occurs before τ_2. For this reason, as in Theorem 1, we define Conditions N1 and A1 to

ensure $|\xi^0_{\tau_2}| \ll N$ (in other words, to ensure stochastic tunneling).

We find it useful to approximate τ_2 with σ_2, the time at which the first type 1 whose lineage will yield a type 2 is born. For this approximation to be valid, we require $\tau_2 - \sigma_2 \ll \sigma_2$, which is ensured by Conditions N2 and A2.

Finally, since we use the fact that mutation is advantageous in our proof of the $r > 1$ case, we need to ensure that mutations are "advantageous enough" to actually produce different behavior on our timescale. We ensure this by placing conditions on the size of $r - 1$ in Theorem 10.

Under any parameter regime that satisfies these conditions (the specifics of which are saved for Chapter 3), we have the following result:

$$P(\tau_2 > t/c) \to \exp(-t)$$

Where c is a constant depending on r, d, and the other model parameters, and is given explicitly in Chapter 3.

In the neutral case, analysis follows a similar strategy in each dimension. First, we show there are constants $a(n)$ that cause the following convergence in distribution for each dimension:

$$\left(\frac{|\xi^0_{T_{n\epsilon}+a(n)t}|}{n} \middle| T_{n\epsilon} < \infty \right) \Rightarrow (X_t | X_0 = \epsilon) \tag{1.2.1}$$

Where $T_{n\epsilon}$ is the first time $|\xi^0_t| = n\epsilon$. Second, using X_t and stochastic calculus, we find the limiting probability that $|\xi^0_{T_{n\epsilon}}|$ gives rise to a type 2 before dying out. We use this probability to find the total rate at which type 2's arise, assuming they

do not come from "small families," instances of ξ_t^0 that will not or haven't yet reached size $n\epsilon$. Finally, we bound the rate at which small families produce type 2's and send $\epsilon \to 0$ to conclude that, in fact, type 2's arise from small families with probability zero.

In the advantageous case, some family of type 1's will never die out with probability 1-1/r. We prove that the first type 2 will emerge from this family as it sweeps to fixation. The entire challenge lies in finding the expected man-hours of cells in families that do die out; this is difficult because the transition rate of $|\xi_t^0|$ depends on the size of the boundary $|\delta\xi_t^0|$. For $d = 1$, the boundary always has size two, so it is easy to consider the embedded discrete time chain. For $d \geq 2$, we bound this expectation by treating the system as a biased voter model and using standard interacting particle system tricks (relabeling, considering the dual process). See 3.3 for complete details.

CHAPTER 2

EVOLUTION OF RESISTANCE AND PROGRESSION TO DISEASE DURING CLONAL EXPANSION OF CANCER

2.1 Introduction

The mathematical investigation of cancer began in the 1950s, when Nordling (1953), Armitage and Doll (1954, 1957), and Fisher (1959) set out to explain the age-dependent incidence curves of human cancers. For a nice survey see Frank (2007). Armitage and Doll (1954) noticed that log-log plots of cancer incidence data are linear for a large number of cancer types; for example, colorectal cancer incidence has a slope of 5.18 in men and 4.97 in women. The authors used this observation to argue that cancer is a multi-stage process and results from the accumulation of multiple genetic alterations in a single cell. The math underlying this hypothesis was very simple. Suppose X_i are independent and have an exponential distribution with rates u_i (i.e., the density function is $u_i e^{-u_i t}$ and the mean is $1/u_i$). Noting that the sum $X_1 + \cdots + X_k$ has a density function that is asymptotically

$$u_1 \cdots u_k \frac{t^{k-1}}{(k-1)!} \quad \text{as } t \to 0, \tag{2.1.1}$$

the authors inferred that the slope of the age-incidence curve was the number of stages minus 1, making colon cancer a six-stage process.

Later on, Knudson (1971) performed a statistical analysis of retinoblastoma, a childhood eye cancer. His study showed that familial cases of retinoblastoma have an earlier age of onset than the sporadic cases that emerge in families without a history of the disease. Based on age incidence curves in the two groups, he hypothesized that two mutagenic events or "hits" are necessary to cause cancer

in the sporadic case, but in individuals with the inherited form of the disease, a single hit is sufficient since one mutation is already present at birth. This study led to the concept of a tumor suppressor gene, i.e., a gene which contributes to tumorigenesis if inactivated in both alleles. See Knudson (2001) for a survey.

Knudson's research led to an explosion of papers on the multi-stage theory of carcinogenesis too numerous to list here. Most studies, like the ones cited in the last two paragraphs, merely fit curves to data on age specific incidence without considering a population genetic model for the cell population. Iwasa et al. (2004,2005) were the first to study waiting times in this way. They used a Moran model for a population of a fixed size N in which type i cells are those with $i \geq 0$ mutations, and type i mutates to type $i+1$ at rate u_{i+1}. Let τ_k be the first time at which there is a type k-cell. They considered a variety of scenarios based on the relative fitnesses of mutants. In the neutral case, i.e., if the mutation does not alter the fitness or growth rate of the cell, they showed:

Theorem 2. *In a population of N cells, τ_2 is approximately exponentially distributed with rate $Nu_1u_2^{1/2}$, provided $1/\sqrt{u_2} \ll N \ll 1/u_1$.*

They called this result "stochastic tunneling" because the 2's arise before the 1's reach fixation. Durrett, Schmidt, and Schweinsberg (2009), see also Schweinsberg (2008), generalized this result to cover τ_k.

In many cases, such as leukemia and polyps in colon cancer, the cell population does not have constant size. For these reasons, Iwasa, Nowak, and Michor (2006) considered the time to develop one mutation in an exponentially growing population and Haeno, Iwasa, and Michor (2007) extended the analysis to waiting for two mutations. Their model is a multi-type branching process in which type i cells are those with $i \geq 0$ mutations. Type-i cells give birth at rate a_i

and die at rate b_i, where $\lambda_i = a_i - b_i > 0$. The previous papers consider a number of different possibilities but here will restrict our attention to the case in which $i \to \lambda_i$ is increasing.

We suppose that during their lifetimes, type-i cells mutate at rate u_{i+1} becoming type $i+1$'s. This is slightly different than the previous approach of having mutations with probability u_{i+1} at birth, which translates into a mutation rate of $a_i u_{i+1}$, and this must be kept in mind when comparing results. In applications, the mutation rates are small compared to birth and death rates, so the reduction of the birth rate of type-i's to $a_i(1-u_{i+1})$ is an insignificant difference.

2.1.1 Growth of type-0's

The number of type-0 cells, $Z_0(t)$, is a branching process, so if $Z_0(0) = 1$, $EZ_0(t) = e^{\lambda_0 t}$ and $e^{-\lambda_0 t}Z_0(t)$ is a nonnegative martingale. Well known results imply that $e^{-\lambda_0 t}Z_0(t) \to W_0$ as $t \to \infty$. A closed-form formula for the generating function $Ex^{Z_0(t)}$ is known, see (2.2.2). To find the Laplace transform of W_0, we let $x = \exp(-\theta e^{-\lambda_0 t})$ in the closed form solution and look at the limit as $t \to \infty$ to conclude

$$Ee^{-\theta W_0} = \frac{b_0}{a_0} + \left(1 - \frac{b_0}{a_0}\right)\frac{1 - b_0/a_0}{1 - b_0/a_0 + \theta}$$

From this we see that if δ_0 is a pointmass at 0, and $\lambda_0 = a_0 - b_0$

$$W_0 =_d \frac{b_0}{a_0}\delta_0 + \frac{\lambda_0}{a_0}\,\text{exponential}(\lambda_0/a_0) \tag{2.1.2}$$

where the exponential(r) distribution has density re^{-rt} and mean $1/r$.

If we let $\Omega_0^0 = \{Z_0(t) = 0 \text{ for some } t \geq 0\}$ then (2.2.1) below implies $P(\Omega_0) = b_0/a_0$, i.e., $W_0 = 0$ if and only if the process dies out. Letting $\Omega_\infty^0 = \{Z_0(t) > 0$ for

all $t \geq 0\}$ we have

$$(e^{-\lambda_0 t} Z_0(t) | \Omega_\infty^0) \to V_0 = \text{ exponential}(\lambda_0/a_0) \tag{2.1.3}$$

and hence the Laplace transform

$$E e^{-\theta V_0} = \frac{\lambda_0}{\lambda_0 + a_0\theta} = (1 + c_{\theta,0}\theta)^{-1} . \tag{2.1.4}$$

where $c_{\theta,0} = a_0/\lambda_0$. Here and in what follows, c's are constants that only depend on the birth and death rates, and not on the mutational rates.

2.1.2 Type-1 Results

Let τ_1 be the time of occurrence of the first type-1. Since type-1's are produced at rate $u_1 Z_0(t)$,

$$P(\tau_1 > t | Z_0(s), s \leq t, \Omega_\infty^0) = \exp\left(-u_1 \int_0^t Z_0(s) ds\right) \tag{2.1.5}$$

τ_1 will occur when $\int_0^t Z_0(s)\, ds$ is of order $1/u_1$. A typical choice for $u_1 = 10^{-5}$, so $1/u_1$ is a large number, and we can use the approximation $(Z_0(s)|\Omega_\infty^0) \approx e^{\lambda_0 s} V_0$. Evaluating the integral, taking the expected value, and using (2.1.4), we conclude that

$$\begin{aligned} P(\tau_1 > t | \Omega_\infty^0) &\approx E \exp\left(-u_1 V_0 (e^{\lambda_0 t} - 1)/\lambda_0)\right) \\ &= \frac{\lambda_0}{\lambda_0 + a_0 u_1 (e^{\lambda_0 t} - 1)/\lambda_0} = \left(1 + c_{\tau,1} u_1 (e^{\lambda_0 t} - 1)\right)^{-1} \end{aligned} \tag{2.1.6}$$

where $c_{\tau,1} = a_0/\lambda_0^2$. The median $t_{1/2}^1$ of the distribution has $\lambda_0^2 = a_0 u_1 (e^{\lambda_0 t_{1/2}^1} - 1)$ so

$$t_{1/2}^1 = \frac{1}{\lambda_0} \log\left(1 + \frac{\lambda_0^2}{a_0 u_1}\right) \tag{2.1.7}$$

Figure 1 shows that (2.1.6) agrees well with the values of τ_1 observed in simulations. Parameters are given in the figure caption.

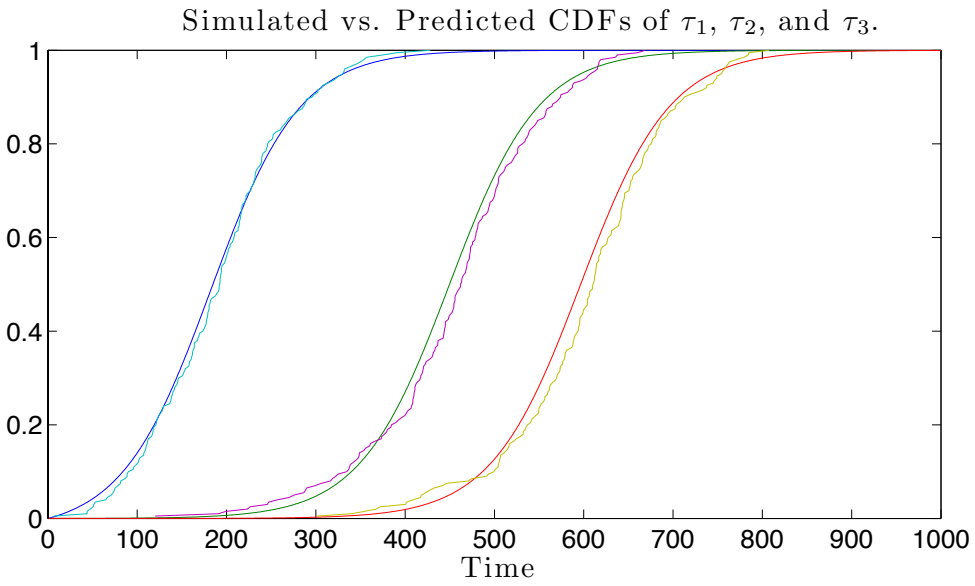

Figure 2.1: Results of 200 runs of the system with $a_0 = 1.02$, $a_1 = 1.04$, $a_2 = 1.06$ $b_i = 1.0$, $u_i = 10^{-5}$. Smooth curves the limit results for τ_i when $i = 1, 2, 3$.

Our next step is to consider the growth of $Z_1(t)$. In Section 2.3 we show that

$$M_t = e^{-\lambda_1 t} Z_1(t) - \int_0^t u_1 e^{-\lambda_1 s} Z_0(s)\, ds \text{ is a martingale}$$

and use this to conclude

Theorem 3. *$e^{-\lambda_1 t} Z_1(t) \to W_1$ a.s. with*

$$EW_1 = u_1/(\lambda_1 - \lambda_0).$$

On Ω_∞^0 we will eventually get a type-1 mutant with an infinite line of descent so $\{W_1 > 0\} \supset \{\Omega_\infty^0\}$.

Let $T_M = \min\{t : Z_0(t) = M\}$. The results of simulations given in Figure 3 of Iwasa, Nowak, and Michor (2006) show that when $\log P(W_1 > x | T_M < \infty)$ is plotted versus $\log x$, a straight line results. Since their M is large, this suggests that $(W_1 | \Omega_\infty^0)$ has a power law tail. As we will now show, this is only approximately

correct. To begin, we consider $Z_i^*(t)$, the number of type-i's at time t in a system with $Z_0^*(t) = e^{\lambda_0 t} V_0$ for all $t \in (-\infty, \infty)$. Let

$$c_{h,1} = \frac{1}{\lambda_0} \left(\frac{a_1}{\lambda_1} \right)^{\lambda_0/\lambda_1 - 1} \Gamma(1 - \lambda_0/\lambda_1)\Gamma(\lambda_0/\lambda_1 + 1)$$

Theorem 4. *$e^{-\lambda_1 t} Z_1^*(t) \to V_1$ a.s. with*

$$E e^{-\theta V_1} = 1/(1 + c_{\theta,1} u_1 \theta^{\lambda_0/\lambda_1})$$

where $c_{\theta,1} = c_{\theta,0} c_{h,1}$, and hence

$$P(V_1 > x) \sim c_{V,1} u_1 x^{-\lambda_0/\lambda_1}$$

where $c_{V,1} = c_{\theta,1}/\Gamma(1 - (\lambda_0/\lambda_1))$.

Iwasa, Nowak, and Michor (2006)'s $\alpha = \lambda_0/\lambda_1$, so our result is consistent with the conclusions given in their (15a) and (15b). The big values of V_1 come from mutations at negative times, so W_1 does not have a power law tail. To upper bound the difference between the distributions of W_1 and V_1 note that the expected number of type-1's produced at times $t \le 0$ is $u_1 a_0/\lambda_0^2$. In the concrete example considered in Figure 1, $a_0 = 1.02$, $b_0 = 1$, and $u = 10^{-5}$ which is 0.0255 so this does not change the limiting distribution by much and the simulated distributions will look like power laws.

A useful consequence of the proof of Theorem 4 is

Corollary. *If we condition on the value of V_0 then $V_1 = \lim_{t\to\infty}$ is the sum of points of a Poisson process on $(0, \infty)$ with intensity $C u_1 V_0 x^{-\lambda_0/\lambda_1}$.*

Here the Poisson points are the sizes of the contributions of different mutations to the limit V_1.

2.1.3 Type-2 Results

We can derive an approximation for the waiting time for the first type 2, τ_2, by using the same reasoning in (2.1.5) and (2.1.6) for τ_1.

$$P(\tau_2 > t | Z_1(s), s \le t, \Omega_\infty^0) \approx \exp\left(-u_2 V_1 e^{\lambda_1 t}/\lambda_1\right) \tag{2.1.8}$$

Taking expected values and using Theorem 4, we obtain

$$P(\tau_2 > t | \Omega_\infty^0) \approx \left(1 + c_{\tau,2}\mu_2 e^{\lambda_0 t}\right)^{-1}$$

where $\mu_2 = u_1 u_2^{\lambda_0/\lambda_1}$, $c_{\tau,2} = c_{\theta,1}\lambda_1^{-\lambda_0/\lambda_1}$, and we have omitted the -1 after $e^{\lambda_0 t}$ because it is not important in this result. Solving we get an approximation for the median value of τ_2:

$$t_{1/2}^2 \approx \frac{1}{\lambda_1}\log\left(\frac{1}{u_2}\right) + \frac{1}{\lambda_0}\log\left(\frac{1}{u_1 c_{\tau,2}}\right) \tag{2.1.9}$$

and it follows easily that

$$P(\tau_2 > t_{1/2}^2 + x/\lambda_0) \to \frac{1}{1+e^x} \tag{2.1.10}$$

Figure 1 compares (2.1.10) with simulations of τ_2.

2.1.4 Type-k Results

To study the growth of the number of type k's for $k \ge 2$, we note that

$$e^{-\lambda_k t} Z_k(t) - \int_0^t u_k e^{-\lambda_k s} Z_{k-1}(s) ds \quad \text{is a martingale}$$

and use this conclude that

Theorem 5. *For $k \ge 2$, $e^{-\lambda_k t} Z_k(t) \to W_k$ a.s. with*

$$EW_k = \prod_{j=1}^{k} \frac{u_j}{\lambda_k - \lambda_{j-1}}$$

Using the approach in the proof of Theorem 4 we can show that if we let

$$c_{h,k} = \frac{1}{\lambda_{k-1}} \left(\frac{a_k}{\lambda_k} \right)^{\lambda_{k-1}/\lambda_k - 1} \Gamma(1 - \lambda_{k-1}/\lambda_k)\Gamma(\lambda_{k-1}/\lambda_k + 1)$$

and $\mu_k = \prod_{j=1}^{k} u_j^{\lambda_0/\lambda_{j-1}}$ then we have

Theorem 6. *$e^{-\lambda_k t} Z_k^*(t) \to V_k$ a.s. with*

$$E e^{-\theta V_k} = \left(1 + c_{\theta,k} \mu_k \theta^{\lambda_0/\lambda_k}\right)^{-1}$$

and hence $P(V_k > x) \sim c_{V,k} \mu_k x^{-\lambda_0/\lambda_k}$, where $c_{V,k} = c_{\theta,k}\Gamma(1 - \lambda_0/\lambda_k)$.

As before, this gives us estimates for the waiting time distribution

$$\begin{aligned} P(\tau_{k+1} > t | \Omega_\infty^0) &\approx E \exp(-V_k u_{k+1} e^{\lambda_k t}/\lambda_k) \\ &= \left(1 + c_{\tau,k+1} \mu_{k+1} e^{\lambda_0 t}\right)^{-1} \end{aligned}$$

where $c_{\tau,k+1} = c_{\theta,k} c_{h,k}^{\lambda_0/\lambda_k}$. Again, we can solve to find the median

$$t_{1/2}^{k+1} = \sum_{j=1}^{k+1} \frac{1}{\lambda_{j-1}} \log\left(\frac{1}{u_j}\right) + \frac{1}{\lambda_0} \log\left(\frac{1}{c_{\tau,k+1}}\right) \tag{2.1.11}$$

and it follows easily that

$$P(\tau_{k+1} > t_{1/2}^{k+1} + x/\lambda_0) \to \frac{1}{1 + e^x} \tag{2.1.12}$$

Note that the shape of the limit distribution is the same as for τ_2 but is translated in time. Figure 1 compares (2.1.12) whne $k = 3$ with simulations of τ_3.

2.1.5 Fixed size results

In Iwasa, Nowak, and Michor (2006) and Haeno, Iwasa, and Michor (2007), the authors consider the system at T_M, the first time at which there are M type-0 cells. With a little more work, we are able to reproduce and extend their results.

$P(\tau_1 < T_M)$

Using the calculation in (2.1.5),

$$P(\tau_1 > T_M | Z_0(s), s \le T_M, \Omega_\infty^0) = \exp\left(-u_1 \int_0^{T_M} Z_0(s)\, ds\right)$$
$$\approx \exp\left(-Mu_1 \int_0^{\infty} e^{-\lambda_0 s}\, ds\right) = \exp\left(-Mu_1/\lambda_0\right) \tag{2.1.13}$$

If we let $\tilde{Z}_1(t) = (Z_1(t)|Z_0(0) = 0, Z_1(0) = 1)$, i.e., the branching process started with no type 0's and one type 1, then similar reasoning shows

$$P(Z_1(T_M) > 0 | Z_0(s), s \le T_M, \Omega_\infty)$$
$$= 1 - \exp\left(-u_1 \int_0^{T_M} Z_0(s) P(\tilde{Z}_1(T_M - s) > 0)\, ds\right)$$

Using $Z_0(s) \approx Me^{-\lambda_0(T_M - s)}$, changing variables $r = T_M - s$, and using (2.2.4) below to evaluate $P(\tilde{Z}_1(T_M - s) > 0)$ the above

$$\approx 1 - \exp\left(-u_1 M \int_0^{(1/\lambda_0)\log M} e^{-\lambda_0 r} \frac{\lambda_1}{a_1 - b_1 e^{-\lambda_1 r}}\, dr\right)$$

where we have stopped the integral when $Z_0(t_M - r) \approx Me^{-\lambda_0 r} = 1$. Changing variables $y = e^{-\lambda_0 r}$, $dy = -\lambda_0 e^{-\lambda_0 r}\, dr$ the integral becomes

$$\frac{1}{\lambda_0} \int_0^1 \frac{\lambda_1}{a_1 - b_1 y^\alpha}\, dy$$

where $\alpha = \lambda_1/\lambda_0$, which agrees with (7) of Iwasa, Nowak, and Michor (2006) once one changes variables $a_0 = r$, $b_0 = d$, $u_1 = ru$. Their derivation of this result is not completely rigorous because they suppose that the number of resistant cells, R_x, produced when $Z_0(t) = x$ are independent, whereas the occupation times $|\{t \le T_M : Z_t(0) = x\}|$ are correlated, but evidently this does not produce a significant error.

$Z_1(T_M)$

Working backward from T_M, assuming deterministic growth of type-0 cells at rate $e^{\lambda_0 s}$, and using a calculation from the proof of Theorem 4, we can show

$$E\exp\left(-\frac{\theta Z_1(T_M)}{(Mu_1)^{\lambda_1/\lambda_0}}\right) \approx \exp\left(-u_1\int_{-\infty}^{0} Me^{\lambda_0 s}(1-\tilde{\phi}_{-s}(\theta(Mu_1)^{-\lambda_1/\lambda_0}))\,ds\right)$$

This leads to

Theorem 7. *As $M\to\infty$, $Z_1(T_M)/(Mu_1)^{\lambda_1/\lambda_0}$ converges to U_1 in distribution where*

$$E(\exp(-\theta U_1)) = \exp(-c_{1,\theta}u_1\theta^{\lambda_0/\lambda_1})$$

and $c_{1,\theta}$ is the constant in Theorem 2.

As in Theorem 4 it follows that $P(U_1 > x) \sim c_{V,1}u_1x^{-\lambda_0/\lambda_1}$. From Theorem 7 we see that if $(Mu_1)^{\lambda_1/\lambda_0} \ll M$, i.e., $M \ll u_1^{-\lambda_1/(\lambda_0-\lambda_1}$ then Haeno, Iwasa, and Michor (2007) are justified in looking at the time when the number of type 0's reaches M rather than when the total population reaches M, see their page 2211. In the concrete example considered in Figure 1, this is $M \ll 10^{2.5}$.

$P(\tau_2 < T_M)$

Using the reasoning for $P(\tau_1 < T_M)$, one can show

$$P(Z_2(T_M) > 0) \approx 1-\exp\left(-\frac{u_1}{\lambda_0}\int_1^M 1-P\left(\tilde{Z}_2\left(\frac{1}{\lambda_0}\log\left(\frac{M}{x}\right)\right)>0\right)\,dx\right)$$

After a change in notation, this is (3) in Haeno, Iwasa, and Michor (2007). To make the connection see their (A3). However, this formula is not very useful, since $P(\tilde{Z}_2(t) > 0)$ is not easy to compute. See their appendix A. One can get a better formula by using Theorem 7 and (2.1.8) to conclude

$$P(\tau_2 < T_M) \approx E\exp(-u_2U_1(Mu_1)^{\lambda_1/\lambda_0}/\lambda_1) = E\exp(-\theta U_1)$$

with $\theta = u_2(Mu_1)^{\lambda_1/\lambda_0}/\lambda_1$. Using the last result with Theorem 7, one can determine the relative proportions of types 0 and 1 at time τ_2. We leave the details to the reader.

2.1.6 Summary

Here, we have derived results for τ_k, the waiting time for the first type k, in a branching process model for an exponentially growing population of cancerous cells. To obtain simple formulas we considered a modification in which $Z_0^*(t) = e^{\lambda_0 t}V_0$ for all $t \in (-\infty, \infty)$. In this case

$$P(\tau_k > t) \approx \left(1 + c_{\tau,k}\mu_k e^{\lambda_0 t}\right)^{-1}$$

where $\mu_k = \prod_{j=1}^{k} u_j^{\lambda_0/\lambda_{j-1}}$ and $c_{\tau,k}$ is an explicit constant that only depends on the birth and death rates.

$$c_{\tau,k} = \frac{a_0}{\lambda_0}\lambda_{k-1}^{-\lambda_0/\lambda_{k-1}} \prod_{i=1}^{k-1}\left[\frac{1}{\lambda_{i-1}}\left(\frac{a_i}{\lambda_i}\right)^{\lambda_0/\lambda_i - 1}\Gamma(1-\lambda_0/\lambda_i)\Gamma(1+\lambda_0/\lambda_i)\right]^{\lambda_0/\lambda_{i-1}}$$

Note that the exponential is $e^{\lambda_0 t}$ for all values of k. Simulations show that despite the fact that various approximations were made in the derivations, the theoretical results agreed well with simulation.

To obtain results for the waiting times via induction, we had to also consider $Z_k^*(t)$, the number of type-k individuals at time t. $e^{-\lambda_k t}Z_k^*(t) \to V_k$ where

$$Ee^{-\theta V_k} = \left(1 + c_{\theta,k}\mu_k\theta^{\lambda_0/\lambda_k}\right)^{-1}$$

Invoking a Tauberian theorem we then concluded that V_k has a power law tail

$$P(V_k > x) \sim c_{V,k}\mu_k x^{-\lambda_0/\lambda_k}$$

confirming simulations of Iwasa, Nowak, and Michor (2006). These results consider the process at a fixed time t, but lead easily to results for the system at time T_M at which there are M type-0 cells, and can be used to obtain results at time S_M when the total tumor size is M.

The remainder of the paper is devoted to proofs. Section 2.2 establishes the branching process results we need. Theorems 1 and 2 are proved in Section 2.3, Theorem 3 in Section 2.4, Theorem 4 in Section 2.5, and Theorem 5 in Section 2.6.

2.2 Branching process results

We begin by computing the extinction probability, ρ. By considering what happened on the first jump

$$\rho = \frac{b_0}{a_0 + b_0} \cdot 1 + \frac{a_0}{a_0 + b_0} \cdot \rho^2$$

Rearranging gives $a_0\rho^2 - (a_0 + b_0)\rho + b_0 = 0$. Since 1 is a root, the quadratic factors as $(\rho - 1)(a_0\rho - b_0) = 0$, and

$$\rho = \begin{cases} b_0/a_0 & \text{if } a_0 > b_0 \\ 1 & \text{if } a_0 \le b_0 \end{cases} \tag{2.2.1}$$

The generating function $F(x,t) = Ex^{Z_0(t)}$ can been computed by solving a differential equation. On page 109 of Athreya and Ney (1972), or in formula (5) of Iwasa, Nowak, and Michor (2006) we find the solution:

$$F(x,t) = \frac{b_0(x-1) - e^{-\lambda_0 t}(a_0 x - b_0)}{a_0(x-1) - e^{-\lambda_0 t}(a_0 x - b_0)} \tag{2.2.2}$$

Subtracting this from 1 gives

$$1 - F(x,t) = \frac{\lambda_0(x-1)}{a_0(x-1) - e^{-\lambda_0 t}(a_0 x - b_0)} \tag{2.2.3}$$

Setting $x = 0$, we have

$$P(Z_0(t) > 0) = 1 - F(0,t) = \frac{\lambda_0}{a_0 - b_0 e^{-\lambda_0 t}} \tag{2.2.4}$$

$e^{-\lambda_0 t} Z_0(t)$ is a nonnegative martingale and converges to a limit W_0, with $EW_0 = 1$ and

$$\{W_0 > 0\} = \{Z_0(t) > 0 \text{ for all } t\} \equiv \Omega_\infty^0$$

To compute the Laplace transform $Ee^{-\theta W_0}$ when $a_0 > b_0$, we set $x = \exp(-\theta e^{-\lambda_0 t})$ in (2.2.2) to get

$$\frac{b_0(\exp(-\theta e^{-\lambda_0 t}) - 1) - e^{-\lambda_0 t}(a_0 \exp(-\theta e^{-\lambda_0 t}) - b_0)}{a_0(\exp(-\theta e^{-\lambda_0 t}) - 1) - e^{-\lambda_0 t}(a_0 \exp(-\theta e^{-\lambda_0 t}) - b_0)}$$

As $t \to \infty$, $e^{-\lambda_0 t} \to 0$, so $\exp(-\theta e^{-\lambda_0 t}) \to 1$, $\exp(-\theta e^{-\lambda_0 t}) - 1 \sim -\theta e^{-\lambda_0 t}$, and the above simplifies to

$$\approx \frac{-b_0\theta e^{-\lambda_0 t} - e^{-\lambda_0 t}\lambda_0}{-a_0\theta e^{-\lambda_0 t} - e^{-\lambda_0 t}\lambda_0} = \frac{b_0\theta + \lambda_0}{a_0\theta + \lambda_0}$$

Dividing top and bottom of this by a_0 and recalling $\lambda_0 = a_0 - b_0$ we have

$$= \frac{(b_0/a_0)\theta + 1 - (b_0/a_0)}{\theta + 1 - (b_0/a_0)} = \frac{b_0}{a_0} + \left(1 - \frac{b_0}{a_0}\right)\frac{1 - (b_0/a_0)}{\theta + 1 - (b_0/a_0)}$$

To invert the Laplace transform, we note that if δ_0 is the point mass at 0 then $p\delta_0 + (1-p)$exponential(ν) has Laplace transform

$$p + (1-p)\frac{\nu}{\nu + \theta} = \frac{p\theta + \nu}{\theta + \nu}$$

so $p = b_0/a_0$, in agreement (2.2.1), and $\nu = 1 - (b_0/a_0)$.

2.3 Growth of the number of type 1's

Our first result is no harder to prove for a general k than it is for $k = 1$, so to avoid repeating the proof later we do it in general now. By considering the times $s \le t$ at which mutations occur and the growth rate of the resulting branching processes of type-k cells,

$$EZ_k(t) = \int_0^t EZ_{k-1}(s)u_k e^{\lambda_k(t-s)}\,ds \tag{2.3.1}$$

Lemma 2.3.1. *$M_t = e^{-\lambda_k t}Z_k(t) - \int_0^t u_k e^{-\lambda_k s}Z_{k-1}(s)\,ds$ is a martingale.*

Proof. Let $\mathcal{F}_t$ be the σ-field generated by $Z_j(s)$ for $0 \le j \le k$ and $s \le t$. Taking differences

$$M_{t+h} - M_t = e^{-\lambda_k(t+h)} Z_k(t+h) - e^{-\lambda_k(t)} Z_k(t) - \int_t^{t+h} u_k e^{-\lambda_k s} Z_{k-1}(s)\, ds$$

Using the expected value formula (2.3.1) we see that

$$E(Z_k(t+h)|\mathcal{F}_t) = e^{\lambda_k h} Z_k(t) + E\left(\int_t^{t+h} u_k Z_{k-1}(s) e^{\lambda_k(t+h-s)}\, ds \middle| \mathcal{F}_t \right)$$

Multiplying by $e^{-\lambda_k(t+h)}$ gives

$$E\left(e^{-\lambda_k(t+h)} Z_k(t+h) - e^{-\lambda_k t} Z_k(t) - \int_t^{t+h} u_k Z_{k-1}(s) e^{-\lambda_k s}\, ds \middle| \mathcal{F}_t \right) = 0$$

The desired result, $E(M_{t+h} - M_t|\mathcal{F}_t) = 0$, follows. □

Proof of Theorem 3. If $\lambda_1 > \lambda_0$ then $I_1 = \int_0^\infty u_1 e^{-\lambda_1 s} Z_0(s)\, ds$ converges and has

$$EI_1 = u_1 \int_0^\infty e^{-(\lambda_1 - \lambda_0)s}\, ds = u_1/(\lambda_1 - \lambda_0)$$

$X_t = -M_t$ is a martingale with $\sup E(X_t^+) \le EI < \infty$, so by the martingale convergence theorem (see e.g., (2.10) in Chapter 4 of Durrett (2005)), X_t converges to a limit X. Since $I_1(t) = \int_0^t u_k e^{-\lambda_k s} Z_0(s)\, ds \to I_1$ as $t \to \infty$, it follows that $e^{-\lambda_1 t} Z_1(t) \to W_1$. The martingale starts at 0 so $Ee^{-\lambda_1 t} Z_1(t) = EI_1(t) \to EI_1$ and it follows from Fatou's lemma that $EW_1 \le EI_1$.

To conclude that $EW_1 = EI_1$, we will show $\sup_t E(e^{-\lambda_1 t} Z_1(t))^2 < \infty$. We will hold off on the proof until we can use induction to address all W_k at once in Section 2.4, see Lemma 2.4.2.

Proof of Theorem 4. To obtain information about the distribution of V_1, recall that $Z_1^*(t)$ is the number of type-1's at time t in the system with $Z_0^*(t) = e^{\lambda_0 t} V_0$ for $t \in (-\infty, \infty)$, let $\tilde{Z}_1(t)$ be the number of 1's at time t in the branching process with $Z_0(0) = 0$, $Z_1(0) = 1$, and let $\tilde{\phi}_{1,t}(\theta) = Ee^{-\theta \tilde{Z}_1(t)}$.

Lemma 2.3.2. $E\left(e^{-\theta Z_1^*(t)}|V_0\right) = \exp\left(-u_1 \int_{-\infty}^t V_0 e^{\lambda_0 s}(1-\tilde{\phi}_{1,t-s}(\theta))\,ds\right)$

Proof. We begin with the corresponding formula in discrete time:

$$\begin{aligned} E\left(e^{-\theta Z_1^*(n)}\Big|\, Z_0(m), m \le n\right) &= \prod_{m=-\infty}^{n-1} \sum_{k_m=0}^{\infty} e^{-u_1 Z_0(m)} \frac{(u_1 Z_0(m))^{k_m}}{k_m!} \tilde{\phi}_{1,n-m-1}(\theta)^{k_m} \\ &= \prod_{m=-\infty}^{n-1} \exp\left(-u_1 Z_0(m)(1-\tilde{\phi}_{1,n-m-1}(\theta))\right) \\ &= \exp\left(-u_1 \sum_{m=-\infty}^{n-1} Z_0(m)(1-\tilde{\phi}_{1,n-m-1}(\theta))\right) \end{aligned}$$

Breaking up the time-axis into intervals of length h and letting $h \to 0$ and using $Z_0^*(s) = \bar{W}_0 e^{\lambda_0 s}$ gives the result in continuous time. □

Replacing θ by $\theta e^{-\lambda_1 t}$ and letting $t \to \infty$

$$E\left(e^{-\theta V_1}|V_0\right) = \lim_{t\to\infty} \exp\left(-u_1 V_0 \int_{-\infty}^t e^{\lambda_0 s}(1-\tilde{\phi}_{1,t-s}(\theta e^{-\lambda_1 t}))\,ds\right) \tag{2.3.2}$$

To calculate the limit, we note that by (2.1.3)

$$\tilde{Z}_1(t-s)e^{-\lambda_1(t-s)} \Rightarrow \frac{b_1}{a_1}\delta_0 + \frac{\lambda_1}{a_1}\,\text{exponential}(\lambda_1/a_1) \tag{2.3.3}$$

so multiplying by $e^{\lambda_1 s}$ and taking the Laplace transform, we have

$$1-\tilde{\phi}_{t-s}(\theta e^{-\lambda_1 t}) \to \frac{\lambda_1}{a_1}\int_0^\infty (1-e^{-\theta x})(\lambda_1/a_1)e^{\lambda_1 s}e^{-xe^{\lambda_1 s}\lambda_1/a_1}dx \tag{2.3.4}$$

Using this in (2.3.2) and interchanging the order of integration

$$E\left(e^{-\theta V_1}|V_0\right) = \exp\left(-u_1 V_0 h(\theta)\right)$$

where

$$h(\theta) = (\lambda_1^2/a_1^2)\int_0^\infty (1-e^{-\theta x})\left[\int_{-\infty}^\infty e^{\lambda_0 s}e^{\lambda_1 s}e^{-xe^{\lambda_1 s}\lambda_1/a_1}ds\right]dx. \tag{2.3.5}$$

Changing variables $u = xe^{\lambda_1 s}\lambda_1/a_1$, $e^{\lambda_1 s}ds = a_1\,du/(\lambda_1^2 x)$ in the inside integral and then $y = \theta x$, $dy = \theta dx$ in the outside integral

$$\begin{aligned} h(\theta) &= \frac{\lambda_1^2}{a_1^2}\int_0^\infty (1-e^{-\theta x})\left[\int_0^\infty \frac{a_1}{x\lambda_1^2}\left(\frac{a_1 u}{\lambda_1 x}\right)^{\lambda_0/\lambda_1} e^{-u}\,du\right]dx \\ &= \frac{1}{a_1}\left(\frac{a_1\theta}{\lambda_1}\right)^{\lambda_0/\lambda_1}\int_0^\infty (1-e^{-y})y^{-\lambda_0/\lambda_1-1}\,dy\int_0^\infty u^{\lambda_0/\lambda_1}e^{-u}\,du \end{aligned} \tag{2.3.6}$$

To make this easier to evaluate we integrate by parts in the first integral to convert it into

$$\frac{\lambda_1}{\lambda_0}\int_0^\infty e^{-y}y^{-\lambda_0/\lambda_1}\,dy$$

and both integrals are values of the Γ function: $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt$.

At this point we have shown

$$h(\theta) = c_{h,1}\theta^{\lambda_0/\lambda_1} \tag{2.3.7}$$

where the constant

$$c_{h,1} = \frac{1}{\lambda_0}\left(\frac{a_1}{\lambda_1}\right)^{\lambda_0/\lambda_1-1}\Gamma(1-\lambda_0/\lambda_1)\Gamma(\lambda_0/\lambda_1+1) \tag{2.3.8}$$

Taking the expected value of $\exp(-u_1V_0h(\theta))$ now, and using (2.1.4) we have

$$E\left(e^{-\theta V_1}\right) = \frac{1}{1+c_{\theta,1}u_1\theta^{\lambda_0/\lambda_1}} \tag{2.3.9}$$

where $c_{\theta,1} = c_{h,1}a_0/\lambda_0$.

To show that V_1 has a power law tail, we note that as $\theta \to 0$,

$$1 - E\left(e^{-\theta V_1}\right) \sim c_{\theta,1}u_1\theta^{\lambda_0/\lambda_1} \tag{2.3.10}$$

and then use a Tauberian theorem from Feller Volume II (pages 442–446). Let

$$\omega(\lambda) = \int_0^\infty e^{-\lambda x}dU(x)$$

Lemma 2.3.3. *If L is slowly varying and U has an ultimately monotone derivative u, then $\omega(\lambda) \sim \lambda^{-\rho} L(1/\lambda)$ if and only if $u(x) \sim x^{\rho-1} L(x)/\Gamma(\rho)$.*

To use this result we note that if $\phi(\theta)$ is the Laplace transform of the probability distribution F, then integrating by parts gives

$$\int_0^\infty e^{-\theta x} dF(x) = (e^{-\theta x})(F(x) - 1)\Big|_0^\infty - \theta \int_0^\infty e^{-\theta x}(1 - F(x))\, dx$$

so we have

$$1 - \phi(\theta) = \theta \int_0^\infty e^{-\theta x}(1 - F(x))\, dx$$

Using (2.3.10), it follows that

$$\frac{1 - E(e^{-\theta V_1})}{\theta} \sim c_{\theta,1} u_1 \theta^{\lambda_0/\lambda_1 - 1}$$

and we conclude

$$P(V_1 > x) \sim c_{V,1} u_1 x^{-\lambda_0/\lambda_1}$$

where $c_{V,1} = c_{\theta,1}/\Gamma(1 - (\lambda_0/\lambda_1))$.

Proof of the Corollary. If S is the sum of Poisson mean λ number of independent random variables with distribution μ then

$$\begin{aligned} Ee^{-\theta S} &= \sum_{k=0}^\infty e^{-\lambda} \frac{\lambda^k}{k!} \left(\int e^{-\theta x} \mu(dx) \right)^k \\ &= \exp\left(-\lambda + \lambda \int e^{-\theta x} \mu(dx) \right) \\ &= \exp\left(- \int (1 - e^{-\theta}) \lambda \mu(dx) \right) \end{aligned}$$

Let $A = Cu_1 V_0$, $\lambda_\epsilon = \int_\epsilon^\infty A x^{-\lambda_0/\lambda_1}\, dx$ and μ_ϵ have density $\lambda_\epsilon^{-1} A x^{-\lambda_0/\lambda_1}$ on (ϵ, ∞). If S_ϵ is the sum of Poisson mean λ_ϵ number of independent random variables with distribution μ_ϵ then

$$Ee^{-\theta S_\epsilon} = \exp\left(- \int_\epsilon^\infty (1 - e^{-\theta}) A x^{-\lambda_0/\lambda_1}\, dx \right)$$

Letting $\epsilon \to 0$ and comparing with (2.3.6) gives the desired result. □

2.4 Proof of Theorem 5

We begin by computing $EZ_k(t)$ using $EZ_k(t) = \int_0^t EZ_{k-1}(s)u_k e^{\lambda_k(t-s)}ds$.

$$EZ_k(t) = u_1u_2\cdots u_k \sum_{j=0}^{k} \frac{e^{\lambda_j t}}{\Gamma_{j,k}} \qquad \text{for } k \geq 1 \tag{2.4.1}$$

where $\Gamma_{j,k} = \prod_{i\leq k, i\neq j}(\lambda_j - \lambda_i)$.

Proof. Let X_j be independent exponential(γ_j), and let p_k is the density function of $X_0 + \cdots + X_k$, which satisfies the recursion

$$p_k(t) = \int_0^t p_{k-1}(s)\gamma_k e^{-\gamma_k(t-s)}\, ds$$

Armitage (1952) has shown, see his paragraph 4, that

$$p_k(t) = (-1)^{k+1}\gamma_0\cdots\gamma_k \sum_{j=0}^{k} \frac{e^{\lambda_j t}}{\Delta_{j,k}}$$

where $\Delta_{j,k} = \prod_{i\leq k, i\neq j}(\gamma_i - \gamma_i)$. If we take $\gamma = -\lambda_i$ then comparing the two recursions and their initial condition $EZ_0(t) = e^{\lambda_0 t}$ and $p_0(t) = \gamma_0 e^{-\gamma_0 t}$ shows

$$p_k(t) = (-1)^{k+1} EZ_k(t)\frac{\lambda_0\cdots\lambda_k}{u_1\cdots u_k}$$

The derivation of the formula for $p_k(t)$ only uses calculus which relies on the γ_i are distinct, so the desired result follows. □

Let $I_k(t) = \int_0^t u_i e^{-\lambda_i s} Z_{k-1}(s)ds$ and $I_k = I_k(\infty)$.

Lemma 2.4.1. *For $k \geq 1$, $EI_k < \infty$.*

Proof Using $EZ_0(t) = e^{\lambda_0 t}$ and (2.4.1)

$$EI_k = E\int_0^\infty u_k e^{-(\lambda_k - \lambda_{k-1})s}\left(e^{-\lambda_{k-1}s}Z_{k-1}(s)\right)ds < \infty$$ □

To prove Theorem 5 now, observe that $X_t = I_k(t) - e^{-\lambda_k t} Z_k(t) \le I_k$ is a martingale and dominated by an integrable random variable, so (2.10) of Chapter 4 of Durrett (2005) implies $X_t \to X$ a.s. Since $I_k(t) \to I_k$ a.s., it follows that $e^{-\lambda_k t} Z_k(t) \to W_k$. (2.4.1) implies that

$$E e^{-\lambda_k t} Z_k(t) \to \frac{u_1 u_2 \cdots u_k}{\Gamma_{k,k}}$$

To prove that $EW_k = EI_k$ we will show

Lemma 2.4.2. *For $k \ge 0$, $\sup_t E(e^{-\lambda_k t} Z_k(t))^2 < \infty$.*

Proof. The base case is easy. We look at the derivative $\frac{d}{dt} E(e^{-\lambda_0 t} Z_0(t))^2$

$$= -2\lambda_0 E(e^{-\lambda_0 t} Z_0(t))^2 + e^{-2\lambda_0 t} \left(E[a_0 Z_0(t)(2Z_0(t) + 1)] - E[b_0 Z_0(t)(2Z_0(t) - 1)] \right)$$

$$= e^{-2\lambda_0 t}(a_0 + b_0) E Z_0(t) = e^{-\lambda_0 t}(a_0 + b_0)$$

And it follows that $\sup_t E(e^{-\lambda_0 t} Z_0(t))^2 < \infty$. Next, we suppose $\sup_t E(e^{-\lambda_{k-1} t} Z_{k-1}(t))^2 \le c_{k-1} < \infty$ and consider the derivative $\frac{d}{dt} E(e^{-\lambda_k t} Z_k(t))^2$

$$\begin{aligned} &= -2\lambda_k E(e^{-\lambda_k t} Z_k(t))^2 + e^{-2\lambda_k t} E[a_k Z_k(t)(2Z_k(t) + 1)] \\ &\quad - e^{-2\lambda_k t} E[b_k Z_k(t)(2Z_k(t) - 1)] + e^{-2\lambda_k t} E[u_k Z_{k-1}(t)(2Z_k(t) + 1)] \\ &= (a_k + b_k) e^{-2\lambda_k t} E Z_k(t) + u_k e^{-2\lambda_k t} E[Z_{k-1}(t)(2Z_k(t) + 1)] \end{aligned}$$

To bound $2u_k e^{-2\lambda_k t} E[Z_{k-1}(t) Z_k(t)]$, we use the Cauchy-Schwarz inequality and $y^{1/2} \le 1 + y$ for $y \ge 0$ to get

$$\begin{aligned} &\le 2u_k e^{-(\lambda_k - \lambda_{k-1})t} E[e^{-2\lambda_{k-1} t} Z_{k-1}^2(t)]^{1/2} E[e^{-2\lambda_k t} Z_k^2(t)]^{1/2} \\ &\le 2u_k e^{-(\lambda_k - \lambda_{k-1})t} c_{k-1}^{1/2} \left(1 + E[e^{-2\lambda_k t} Z_k^2(t)] \right) \end{aligned}$$

Comparison theorems for differential equations imply that $E(e^{-\lambda_k t} Z_k(t))^2 \le m(t)$ where $m(t)$ is the solution of the differential equation

$$\frac{d}{dt} m(t) = a(t) m(t) + b(t), \quad m(0) = 0 \tag{2.4.2}$$

with $a(t) = 2u_k c_{k-1}^{1/2} e^{-(\lambda_k - \lambda_{k-1})t}$, and

$$b(t) = (a_k + b_k)e^{-2\lambda_k t} EZ_k(t) + 2u_k e^{-2\lambda_k t} EZ_{k-1}(t) + 2u_k c_{k-1}^{1/2} e^{-(\lambda_k - \lambda_{k-1})t}$$

Solving (2.4.2) gives

$$m(t) = \int_0^t b(s) \exp\left(\int_s^t a(r)\, dr \right)$$

Since $a(t)$ and $b(t)$ are both integrable, $m(t)$ is bounded. □

2.5 Proof of Theorem 6

Let $\mathcal{F}_t^{k-1}$ be the σ-field generated by $Z_j^*(s)$ for $j \le k-1$ and $s \le t$. Let $\tilde{Z}_k(t)$ be the number of type k's at time t in the branching process with $Z_k(0) = 1$ and $Z_j(0) = 0$ for $j \le k-1$, and let $\tilde{\phi}_{k,t}(\theta) = Ee^{-\theta \tilde{Z}_1(t)}$. The reasoning of Lemma 2.3.2 implies

$$E(e^{-\theta Z_k^*(t)} | \mathcal{F}_t^{k-1}) = \exp\left(-u_k \int_{-\infty}^t Z_{k-1}^*(s)(1 - \tilde{\phi}_{k,t-s}(\theta))\, ds \right)$$

Replacing $Z_{k-1}^*(s)$ by $e^{\lambda_{k-1} s} V_{k-1}$, θ by $\theta e^{-\lambda_k t}$, and letting $t \to \infty$

$$E\left(e^{-\theta V_k} | \mathcal{F}_\infty^{k-1} \right) = \lim_{t \to \infty} \exp\left(-u_k V_{k-1} \int_{-\infty}^t e^{\lambda_{k-1} s} (1 - \tilde{\phi}_{k,t-s}(\theta e^{-\lambda_k t}))\, ds \right) \tag{2.5.1}$$

At this point the calculation is the same as the one in Section 2.3 with 1 and 0 replaced by k and $k-1$ respectively, and we conclude that

$$E\left(e^{-\theta V_k} | \mathcal{F}_\infty^{k-1} \right) = \exp\left(-u_k V_{k-1} h_k(\theta) \right) \tag{2.5.2}$$

where $h_k(\theta) = c_{h,k} \theta^{\lambda_{k-1}/\lambda_k}$ and

$$c_{h,k} = \frac{1}{\lambda_{k-1}} \left(\frac{a_k}{\lambda_k} \right)^{\lambda_{k-1}/\lambda_k - 1} \Gamma(1 - \lambda_{k-1}/\lambda_k) \Gamma(\lambda_{k-1}/\lambda_k + 1)$$

Let $c_{\theta,k} = c_{\theta,k-1} c_{h,k}^{\lambda_0/\lambda_k}$. When $k = 2$ taking expected value and using Theorem 4 gives

$$Ee^{-\theta V_2} = \left(1 + c_{\theta,2} u_1 u_2^{\lambda_0/\lambda_1} \theta^{\lambda_0/\lambda_2} \right)^{-1}$$

Using this in (2.5.2)

$$Ee^{-\theta V_3} = \left(1 + c_{\theta,3} u_1 u_2^{\lambda_0/\lambda_1} u_3^{\lambda_0/\lambda_2} \theta^{\lambda_0/\lambda_3}\right)^{-1}$$

The pattern should be clear so we leave to the reader to check the induction step. The result for $P(V_k > x)$ follows from Lemma 2.3.3, and the proof of Theorem 6 is complete.

2.6 Proof of Theorem 7

We are interested in finding

$$\lim_{M\to\infty} \exp\left[-u_1 \int_{-\infty}^{0} Me^{\lambda_0 s}(1 - \tilde{\phi}_{-s}(\theta(Mu_1)^{-\lambda_1/\lambda_0}))\, ds\right]$$

First, we make the change of variables $s = t - \frac{1}{\lambda_0}\log(Mu_1)$.

$$= \lim_{M\to\infty} \exp\left[-\int_{-\infty}^{\frac{1}{\lambda_0}\log(Mu_1)} e^{\lambda_0 t}(1 - \tilde{\phi}_{\frac{1}{\lambda_0}\log(Mu_1)-t}(\theta(Mu_1)^{-\lambda_1/\lambda_0}))\, dt\right]$$

Taking the limit as $M \to \infty$ is essentially the same calculation as (2.3.4).

$$= \exp\left[-\int_{-\infty}^{\infty} e^{\lambda_0 t}\frac{\lambda_1}{a_1}\int_0^{\infty}(1 - e^{-\theta x})(\lambda_1/a_1)e^{\lambda_1 t}e^{-xe^{\lambda_1 t}\lambda_1/a_1}\, dx\, dt\right]$$

We conclude by recognizing this double integral as $h(\theta)$ defined in (2.3.5) and computed in (2.3.7).

2.7 References

Armitage, P. (1953) A note on the time-homogeneous birth process. *J. Royal Statist. Soc.* 15, 90–91

Armitage, P., and Doll, R. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Brit. J. Cancer.* 8, 1–12

Armitage, P., and Doll, R. (1957) A two-stage theory of carcinogenesis. *Brit. J. Cancer.* 11, 161–169

Durrett, R. (2005) *Probability: Theory and Examples.* Third Edition, Duxbury Press

Durrett, R., Schmidt, D., and Schweinsberg, J. (2009). A waiting time problem arising from the study of multi-stage carcinogenesis. *Ann. Appl. Prob.* 19, 676–718.

Fisher, J.C. (1959) Multiple mutation theory of carcinogenesis. *Nature.* 181, 651–652.

Frank, S.A. (2007) *Dynamics of Cancer: Incidence, Inheritance and Evolution.* Princeton Series in Evolutionary Biology

Haeno, H., Iwasa, Y., and Michor, F. (2007) The evolution of two mutations during clonal expansion. *Genetics.* 177, 2209–2221

Iwasa, Y., Michor, F. and Nowak, M.A. (2004) Stochastic tunnels in evolutionary dynamics. Genetics. 166, 1571–1579

Iwasa, Y., Michor, F., Komarova, N.L. and Nowak, M.A. (2005) Population genetics of tumor suppressor genes.*J. Theor. Biol.* 233, 15–23

Iwasa, Y., Nowak, M.A., and Michor, F. (2006) Evolution of resistance during clonal expansion. *Genetics.* 172, 2557–2566

Knudson, A.G. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. USA* 68, 820–823

Knudson, A.G. (2001) Two genetic hits (more or less) to cancer. *Nature Reviews Cancer.* 1, 157–162

Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., et al. (2005) Dynamics of chronic myeloid leukemia. *Nature.* 435, 1267–1270

Nordling, C.O. (1953) A new theory on cancer inducing mechanisms. *British J. Cancer.* 7, 68–72

Schweinsberg, J. (2008) The waiting time for m mutations. *Electron. J. Probab.* 13, 1442–1478.

CHAPTER 3

A SPATIAL MODEL FOR TUMOR GROWTH

3.1 Introduction

The mathematical investigation of cancer began in the 1950s, when Nordling (1953), Armitage and Doll (1954, 1957), and Fisher (1959) set out to explain the age-dependent incidence curves of human cancers. For a nice survey, see Frank (2007). Armitage and Doll (1954) noticed that log-log plots of cancer incidence data are linear for a large number of cancer types; for example, colorectal cancer incidence has a slope of 5.18 in men and 4.97 in women. The authors used this observation to argue that cancer is a multi-stage process and results from the accumulation of multiple genetic alterations in a single cell. Thinking about sums of independent exponentially distributed random variables, the authors inferred that the slope of the age-incidence curve was the number of stages minus 1, making colon cancer a six-stage process.

Later on, Knudson (1971) performed a statistical analysis of retinoblastoma, a childhood eye cancer. His study showed that familial cases of retinoblastoma have an earlier age of onset than the sporadic cases that emerge in families without a history of the disease. Based on age incidence curves in the two groups, he hypothesized that two mutagenic events or "hits" are necessary to cause cancer in the sporadic case, but in individuals with the inherited form of the disease, a single hit is sufficient since one mutation is already present at birth. This study led to the concept of a tumor suppressor gene, i.e., a gene which contributes to tumorigenesis if inactivated in both alleles. See Knudson (2001) for a survey.

The accumulation of several mutations is important not only for cancer initiation and progression, but also for the emergence of acquired resistance against chemotherapeutics, radiation therapy, or targeted drugs. For this reason there is a large and growing literature on the waiting time τ_k until some cell has acquired k prespecified mutations. In all the models we will discuss type i individuals are cells with i mutations, and type i individuals mutate to type $(i+1)$ at rate u_{i+1}, which we assume to be much less than 1. The dynamics considered are primarily of two types: (i) multi-type branching processes, and (ii) multi-type Moran models.

Here we will concentrate on models of the second type and focus on τ_2. Throughout this paper, we suppose that cells of type 0 and type 1 have relative fitness 1 and r. Since we will only consider the waiting time for the first type 2, the relative fitness of type 2's is not important.

3.1.1 Neutral mutations

Komarova, Sengupta and Nowak (2003) proved the following:

Theorem 8. *In the neutral case, $r = 1$, if we assume that*

$$\frac{1}{\sqrt{u_2}} \ll N \ll \frac{1}{u_1} \tag{3.1.1}$$

then we have $P(\tau_2 > t/Nu_1u_2^{1/2}) \to \exp(-t)$.

See also Iwasa, Michor, and Nowak (2004), Iwasa et al. (2005), and Wodarz and Komarova (2005). This result is called "stochastic tunneling" because the 2's arise before the 1's reach fixation. For an intuitive discussion, see Section 12.4 of Nowak's (2006) book. Durrett and Schmidt (2008) applied these ideas

to study regulatory sequence evolution and to expose flaws in Michael Behe's arguments for intelligent design. Durrett, Schmidt, and Schweinsberg (2009), see also Schweinsberg (2008), generalized this result to cover τ_k.

For simplicity, many multistage cancer models assume a homogeneously mixing cell populations. However, this is not realistic for many solid tumors. For this reason, Komarova (2006) considered a spatial model, which had been introduced much earlier by Williams and Bjerknes (1972). We begin with the version with an infinite number of cells, since it is easier to describe: each location on the d-dimensional integer lattice $\mathbb{Z}^d$ is inhabited by a single cell, initially all of type 0. At times of a Poisson process with rate 1, each cell x selects one of its $2d$ nearest neighbors y, chosen uniformly at random. If the two cells are the same type, nothing happens, but if they differ, x adopts the type of y with probability $1/(r+1)$ if y is type 0, and $r/(1+r)$ if y is type 1.

When there are no mutations and $r = 1$ this is the voter model which was introduced independently by Clifford and Sudbury and Holley and Liggett (1975). For a summary of what is known see Liggett (1999). Since we want a finite cell population we will restrict our process to be a subset of $[-L/2, L/2)^d$ and denote it by $\bar{\xi}_t^0$. Komarova (2006) uses "Dirichlet boundary conditions", i.e., she assumes her space is an interval with no cells outside, but this is awkward because the set of type 1 cells may reach one end of the interval and then no further changes happen at that end. To avoid this, we will use periodic boundary conditions, i.e., we consider $(\mathbb{Z} \bmod L)^d$. When $d = 2$ this means we are thinking of your skin as a torus, which is a little odd. However, using $(\mathbb{Z} \bmod L)^d$ has the advantage that the space looks the same seen from any point. Our results will show that (for the parameter values we consider) the first type 2 will arise when

the radius of $\bar{\xi}_t^0$ is $\ll L$ so the boundary conditions do not matter.

Let ξ_s^0 be the set of cells equal to 1 in the voter model on $\mathbb{Z}^d$ starting from a single type 1 at 0, let $|\xi_s^0|$ be the number of cells in ξ_s^0, and let

$$\nu_d = 1 - E\exp\left(-u_2 \int_0^{T_0} |\xi_s^0|\, ds\right) \tag{3.1.2}$$

be the probability, which depends on the dimension d, that a mutation to type 1 gives rise to a type 2 before its family dies out, assuming there were no further mutations to type 1. Since mutations to type 1 in a population of N cells occur at rate Nu_1 this suggests that

$$P(\tau_2 > t) \approx \exp\left(-Nu_1\nu_d t\right) \tag{3.1.3}$$

As we will explain in a moment, $\nu_d \sim \gamma_d h_d(u_2)$ as $u_2 \to 0$ where

$$h_d(u) = \begin{cases} u^{1/3} & d = 1 \\ u^{1/2}\log^{1/2}(1/u) & d = 2 \\ u^{1/2} & d \geq 3 \end{cases} \tag{3.1.4}$$

and $f(u) \sim g(u)$ means $f(u)/g(u) \to 1$. To state the result we need one more definition:

$$g_d(u) = \begin{cases} u^{1/3} & d = 1 \\ \log^{-1/2}(1/u) & d = 2 \\ 1 & d \geq 3 \end{cases} \tag{3.1.5}$$

Theorem 9. *In the neutral case, $r = 1$, if we assume*

$$\frac{1}{h_d(u_2)} \ll N \ll \frac{g_d(u_2)}{u_1} \tag{3.1.6}$$

then there are constants γ_d given in (3.1.8) and (3.1.9) so that

$$P(\tau_2 > t/Nu_1\gamma_d h_d(u_2)) \to \exp(-t)$$

In $d = 1$ this result was proved by Komarova (2006), see her equation (62) and assumption (60), then change notation $u_1 \to u$, $u_2 \to u_1$. Note that when $d \geq 3$ the order of magnitude of the waiting time and the assumptions are the same as in Theorem 8. In $d = 2$ there are logarithmic corrections to the behavior in Theorem 8, so only in the biologically unrealistic case of $d = 1$ does space make a substantial change in the waiting time.

The reasons for the two conditions in Theorem 9 are the same as in Theorem 8.

(N1) The left hand assumption in (3.1.6) implies that $|\xi^0_{\tau_2}| \ll N$, so the type 2 mutant arises before the 1's reach fixation.

(N2) Let σ_2 be the time of the first type 1 mutation that leads to a type 2. Since mutations to type 1 occur at rate Nu_1 and lead to a type 2 with probability ν_d, it is easy to see that

$$P(\sigma_2 > t) \approx \exp(-Nu_1\nu_d t)$$

so to prove the result we need to show that with high probability $\tau_2 - \sigma_2 \ll \sigma_2$, and this is guaranteed by the right-hand assumption.

The next order of business is to explain how (N1) and (N2) translate into (3.1.6), derive (3.1.4) and identify the constants. Let T_k be the first time $|\xi^0_t| = k$. The key is to the proof of Theorem 9 is to show that there are constants $a(n)$

$$\left(\frac{|\xi^0_{T_{n\epsilon}+a(n)t}|}{n} \middle| T_{n\epsilon} < \infty \right) \Rightarrow (X_t | X_0 = \epsilon) \tag{3.1.7}$$

where $\Rightarrow$ indicates convergence in distribution of the stochastic processes. In $d = 1$, $|\xi^0_t|$ is a simple random walk that jumps at rate 2 and has an absorbing state at 0, so if we take $a(n) = n^2$ then X_t is a Brownian motion with variance 2t killed when it hits 0.

The one dimensional case is easy because the boundary of ξ_t^0 always has size 2 when $|\xi_t^0| > 0$. In $d \geq 2$ when $|\xi_t^0| = k$ the boundary could be as large as $2dk$ for isolated points, or as small as $Ck^{1/d}$ for a solid ball. The right answer turns out to be in between these two extremes, but much closer to the upper bound. Let $\beta_2 = \pi$ and for $d \geq 3$ let β_d be the probability two simple random walks started at 0 and $e_1 = (1, 0, \ldots 0)$ never hit. Sawyer (1979) and Bramson and Griffeath (1980a) were the first to study the size of $|\xi_t^0|$, but here we need the more sophisticated result of Bramson, Cox, and LeGall (2001). They have shown that if we take $a(n) = n \log n$ in $d = 2$ and $a(n) = n$ in $d \geq 3$ then (3.1.7) holds with X_t = the Feller's branching diffusion, which solves the SDE

$$dX_t = \sqrt{2\beta_d X_t}\, dW_t$$

As we will explain in detail later, this limit results because when $|\xi_t^0| = k$ and k is large then the size of the boundary is $\sim 2d\beta_d k$ in $d \geq 3$ and $\sim 4\beta_2 k/\log k$ in $d = 2$.

Let ν_d^ϵ be the probability defined in (3.1.2) ignoring mutations to type 2 that occur before $T_{n\epsilon}$. $|\xi_t^0|$ is a martingale, so $P_1(T_{n\epsilon} < \infty) = 1/n\epsilon$ and using (3.1.7)

$$\nu_d^\epsilon \approx \frac{1}{n\epsilon} \cdot \left[1 - E_\epsilon \exp\left(-na(n)u_2 \int_0^{T_0} X_s\, ds \right) \right]$$

where $T_0 = \min\{t : X_t = 0\}$. $na(n) = n^3$ in $d = 1$, $n^2 \log n$ in $d = 2$, and n^2 in $d \geq 3$, so if we set $n = 1/h_d(u_2)$ then

$$\nu_d^\epsilon \approx h_d(u_2) \cdot \left[\frac{1 - E_\epsilon \exp\left(-\int_0^{T_0} X_s\, ds \right)}{\epsilon} \right]$$

Stochastic calculus (or calculations with infinitesimal generators) tells us that

$$v(x) = E_x \exp\left(-\int_0^{T_0} X_s\, ds \right)$$

is the unique function on $[0, \infty)$ with values in $[0, 1]$, $v(0) = 1$ and

$$v'' - xv = 0 \quad \text{in } d = 1 \qquad \beta_d x v'' - xv = 0 \quad \text{in } d \geq 2$$

In $d = 1$ all solutions have the form:

$$v(x) = \alpha Ai(x) + \beta Bi(x)$$

where Ai and Bi are Airy functions

$$Ai(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{t^3}{3} + xt\right) dt$$
$$Bi(x) = \frac{1}{\pi} \int_0^\infty \exp\left(-\frac{t^3}{3} + xt\right) + \sin\left(\frac{t^3}{3} + xt\right) dt.$$

Since Bi is unbounded and Ai is decreasing on $[0, \infty)$, we take $\beta = 0$ and set $\alpha = 3^{2/3}\Gamma(2/3)$ to satisfy the boundary condition, $v(0) = 1$. Letting $\epsilon \to 0$ we conclude that

$$\gamma_1 = -\alpha Ai'(0) = 3^{1/3}\Gamma(2/3)/\Gamma(1/3) \tag{3.1.8}$$

In $d \geq 2$, $v(x) = \exp(-\beta_d^{-1/2} x)$, and we have

$$\gamma_d = \beta_d^{-1/2} \tag{3.1.9}$$

Looking back at the calculations above, we see that the mutation to type 2 is likely to occur in a type 1 family that reaches size $1/h_d(u_2)$ so for (N1) to hold we need $1/h_d(u_2) \ll N$. The time needed to reach this level is, by (3.1.7),

$$a(1/h_d(u_2)) = \begin{cases} u_2^{-2/3} & d = 1 \\ u_2^{-1/2} \log^{1/2}(1/u_2) & d = 2 \\ u_2^{-1/2} & d \geq 3 \end{cases}$$

Thus for (N2) we need $a(1/h_d(u_2)) \ll 1/Nu_1 h_d(u_2)$, which means $N \ll g_d(u_2)/u_1$

3.1.2 Advantageous mutations

We turn now to the case in which type 1 cells have fitness $r > 1$ relative to type 0 cells. In any dimension $|\xi_t^0|$ is a time change of simple random walk that jumps up with probability $r/(r+1)$ and down with probability $1/(r+1)$, so well known results imply that

$$P_1(T_0 = \infty) = 1 - \frac{1}{r}$$

By analogy with (3.1.3) we expect that

$$P(\tau_2 > t) \approx \exp(-tNu_1(r-1)/r)$$

For this approximation to be accurate we need some assumptions. The first two are the analogues of (N1) and (N2).

(A1) Results of Bramson and Griffeath (1980b, 1981) imply that when ξ_t^0 does not die out it grows linearly and has an asymptotic shape. Thus if the type 2 mutation occurs before type 1's reach fixation, it will occur when

$$\int_0^{\tau_2-\sigma_2} s^d \, ds = O(1/u_2).$$

That is, $\tau_2 - \sigma_2 = O(u_2^{-1/(d+1)})$. Since at that time the number of type 1's is $O(u_2^{-d/(d+1)})$ we need to have $u_2^{-d/(d+1)} \ll N$.

(A2) As before, it is easy to see that

$$P(\sigma_2 > t) \approx \exp(-tNu_1(r-1)/r)$$

so to prove the result we need to show that with high probability $\tau_2 - \sigma_2 \ll \sigma_2$ which requires

$$u_2^{-1/(d+1)} \ll 1/Nu_1(r-1)$$

We have dropped the factor of r because we are only concerned with $1 < r \le R$ where R is fixed.

(A3) Finally, we have to consider the possibility that the mutation to type 2 occurs among the descendants of a type 1 mutation that does not reach fixation. In the neutral case this does not require an additional condition but here it does.

Theorem 10. *In the advantageous case $r > 1$, if we assume that*

$$\frac{1}{u_2^{d/(d+1)}} \ll N \ll \frac{u_2^{1/(d+1)}}{u_1(r-1)} \tag{3.1.10}$$

and $r - 1 \gg h_d(u_2)$ then $P(\tau_2 > tr/Nu_1(r-1)) \to \exp(-t)$.

The first set of conditions is similar to (3.1.1) and (3.1.6). The new one says that $r - 1$ is large enough so that the type 1 mutations are not neutral. To explain this, we note that from the explanation of the proof of Theorem 9, in the neutral case the mutation to type 2 will occur when a family of type 1's reaches size $O(1/h_d(u_2))$. Thus in order for the selective advantage of type 1 to change the behavior from the neutral case, we must have $r - 1 \gg h_d(u_2)$. Komarova (2006) observed that this was the necessary condition for non-neutrality in $d = 1$. Iwasa, Nowak, and Michor (2004) showed that in the homogeneously mixing case the condition is $r - 1 \gg u_2^{1/2}$.

3.2 Proof of Theorem 9

In the introduction we have calculated the probability ν_d^ϵ that a type 1 family reaches size $\epsilon/h_d(u_2)$ and then gives rise to a type 2. To let $\epsilon \to 0$ and prove Theorem 9 we need to consider the possibility of a mutation to type 2 in a family that (i) never reaches size $n\epsilon$, or (ii) will reach $n\epsilon$ but hasn't yet. To have a convenient name we will call these small families. Families of the first kind arise at rate $Nu_1(1 - 1/n\epsilon)$ and families of the second kind arise at rate $Nu_1/n\epsilon$. We will now calculate the expected rate at which type-2's are born from these small families.

Consider the voter model ξ_t^0 starting from a single 1 at the origin at time 0. Let V_k be the total time spent at level k, i.e., $|\{t : |\xi_t^0| = k\}|$, let N_k be the total number of returns to level k before leaving the interval $(0, n\epsilon)$, and let $q(k)$ the rate jumps occur at level k. As the reader will see, $q(k)$ is the only element that depends on dimension.

$$
\begin{aligned}
E_1\left(\int_0^{T_0} |\xi_s^0|\, ds \,\middle|\, T_0 < T_{n\epsilon}\right) &= E_1\left(\sum_{k=1}^{n\epsilon} kV_k \,\middle|\, T_0 < T_{n\epsilon}\right) \\
= E_1\left(\sum_{k=1}^{n\epsilon} \frac{kN_k}{q(k)} \,\middle|\, T_0 < T_{n\epsilon}\right) &= \sum_{k=1}^{n\epsilon} \frac{\bar{P}_1(T_k < \infty)}{\bar{P}_k(T_k^+ > T_0)} \frac{k}{q(k)}
\end{aligned} \tag{3.2.1}
$$

Where the bar indicates conditioning on $T_0 < T_{n\epsilon}$. A similar argument shows that

$$
E_1\left(\int_0^{T_{n\epsilon}} |\xi_s^0|\, ds \,\middle|\, T_{n\epsilon} < T_0\right) = \sum_{k=1}^{n\epsilon} \frac{1}{\hat{P}_k(T_k^+ > T_{n\epsilon})} \frac{k}{q(k)} \tag{3.2.2}
$$

where the hat indicates conditioning on $T_{n\epsilon} < T_0$.

The three conditional probabilities we need can be computed using facts

about simple random walk that follow from the fact that it is a martingale.

$$\bar{P}_1(T_k < \infty) = \frac{P_1(T_k < \infty)P_k(T_0 < T_{n\epsilon})}{P_1(T_0 < T_{n\epsilon})} = \frac{(1/k)(1 - k/n\epsilon)}{(1 - 1/n\epsilon)} \tag{3.2.3}$$

For the next two we note that the first step has to be in the correct direction for these events to happen.

$$\bar{P}_k(T_k^+ > T_0) = \frac{(1/2)(1/k)}{(1 - k/n\epsilon)} \tag{3.2.4}$$

$$\hat{P}_k(T_k^+ > T_{n\epsilon}) = \frac{(1/2)(\frac{1}{n\epsilon - k})}{(k/n\epsilon)} \tag{3.2.5}$$

Thus the expected total man-hours $\int_0^{T_0} |\xi_s^0|\, ds$ for a family that will die out before reaching size $n\epsilon$ is

$$\frac{2}{(1 - 1/n\epsilon)} \sum_{k=1}^{n\epsilon} (1 - k/n\epsilon)^2 \frac{k}{q(k)} \tag{3.2.6}$$

And in families that have yet to reach size $n\epsilon$

$$\frac{2}{n\epsilon} \sum_{k=1}^{n\epsilon} (n\epsilon - k) \frac{k^2}{q(k)} \tag{3.2.7}$$

One Dimension. In one dimension, $q(k) = 2$. The sum in (3.2.6) is dominated by

$$\int_0^{n\epsilon} (1 - x/n\epsilon)^2 x\, dx = \frac{1}{(n\epsilon)^2} \int_0^{n\epsilon} y^2 (n\epsilon - y)\, dy = \frac{(n\epsilon)^2}{12}.$$

Thus, families of the first kind produce type 2's at rate $\leq Nu_1u_2(n\epsilon)^2/12$. The expression in (3.2.7) is dominated by

$$\frac{2}{n\epsilon} \int_0^{n\epsilon} (n\epsilon - x)x^2\, dx = \frac{(n\epsilon)^3}{6}.$$

Thus, families of the second kind produce type-2's at rate $\leq Nu_1u_2(n\epsilon)^2/6$. Type 2's emerge from the process as a whole at rate $O(Nu_1u_2^{1/3})$. Since $n = u_2^{-1/3}$, the rate from small families is $Nu_1u_2^{1/3}\epsilon^2/4$, so their contribution is indeed negligible. Here and in the next two calculations the order of magnitude of the contributions from the two kinds of small families is the same as the overall rate.

Three or more Dimensions. In dimensions $d \geq 3$, the rate of jumps when $|\xi_t^0| = k$ depends on the set ξ_t^0. However, Cox, Durrett, and Perkins (2000) have shown that when k is large the jump rate is close to $2d\beta_d k$ with high probability, see (I1) on page 202 and consult the definitions on pages 186 and 196. The intuition behind this result is that the voter model is dual to a collection of coalescing random walks, so neighbors of points in ξ_t^0 will be occupied with probability $\approx \beta_d$, the probability that the genealogies of the two sites never coalesce. Using $q(k) = 2d\beta_d k$, (3.2.6) becomes

$$\frac{1}{d\beta_d(1 - 1/n\epsilon)} \sum_{k=1}^{n\epsilon} (1 - k/n\epsilon)^2$$

The sum is bounded above by the integral

$$\int_0^{n\epsilon} (1 - x/n\epsilon)^2 \, dx = \frac{n\epsilon}{3},$$

so with our choice of $n = u_2^{-1/2}$, families of the first kind produce type 2's at rate bounded above by $Nu_1 u_2^{1/2}\epsilon/(3d\beta_d)$. Using $q(k) = 2d\beta_d k$, (3.2.7) becomes

$$\frac{1}{d\beta_d n\epsilon} \sum_{k=1}^{n\epsilon} (n\epsilon - k)k$$

The sum is bounded above by the integral

$$\int_0^{n\epsilon} (n\epsilon - x)x \, dx = \frac{(n\epsilon)^3}{6}.$$

Thus, families of the second kind produce type-2's at rate $\leq Nu_1 u_2^{1/2}\epsilon/(6d\beta_d)$. Comparing to the total rate at which type 2's are produced $O(Nu_1 u_2^{1/2})$, we conclude that the total contribution from small families can be ignored.

Two Dimensions. In two dimensions, the recurrence of random walks implies that when $|\xi_t^0| = k$ is large neighbors of points in ξ_t^0 will be occupied with probability close to 1. In this case the result of Cox, Durrett, and Perkins (2000) implies

that when k is large $q(k) \approx 4\beta_2 k/\log k$ with high probability, and (3.2.6) becomes

$$\frac{1}{2\beta_2(1-1/n\epsilon)} \sum_{k=1}^{n\epsilon} (1-k/n\epsilon)^2 \log k$$

Each term in the sum is bounded above by $\log(n\epsilon)$, so the sum is less than $n\epsilon \log n\epsilon$. Since $n = u_2^{-1/2} \log^{-1/2}(1/u_2)$, families of the first kind produce type 2's at rate bounded above by

$$\begin{aligned} Nu_1u_2 \cdot \frac{1}{2\beta_2} n\epsilon \log(n\epsilon) &= Nu_1u_2 \cdot \frac{1}{2\beta_2} \epsilon u_2^{-1/2} \log^{-1/2}(1/u_2) \cdot \frac{1}{2} \log(1/u_2) \\ &= \frac{\epsilon}{4\beta_2} Nu_1 u_2^{1/2} \log^{1/2}(1/u_2) \end{aligned}$$

Taking $q(k) \approx 4\beta_2 k/\log k$, (3.2.7) becomes

$$\frac{1}{2\beta_2 n\epsilon} \left(\sum_{k=1}^{n\epsilon} (n\epsilon - k) k \log k \right)$$

The sum is bounded above by

$$\int_0^{n\epsilon} (n\epsilon - x) x \log(n\epsilon)\, dx \leq \frac{(n\epsilon)^3}{6} \log(n\epsilon)$$

Thus families of the second kind produce type 2's at rate bounded above by

$$\frac{Nu_1u_2}{n\epsilon} \cdot \frac{1}{2\beta_2 n\epsilon} \cdot \frac{(n\epsilon)^3}{6} \log(n\epsilon) = \frac{1}{12\beta_2} Nu_1u_2 \cdot n\epsilon \log(n\epsilon)$$

which is, up to a constant, the same rate for families of the first kind. Comparing with the total rate at which type 2's are produced, $O(Nu_1u_2^{1/2} \log^{1/2}(1/u_2))$, we conclude that the total contribution from small families can again be neglected.

3.3 Proof of Theorem 10

The remaining detail is to estimate the probability that the mutation to type 2 occurs in a family that later dies out. If $\xi_t^0 = A$ with $|A| = k$ and $|\partial A| = \ell$ then $|\xi_t^0|$ grows to size $k+1$ at rate $r\ell$, and shrinks to size $k-1$ at rate ℓ, so the transition probability of the embedded discrete time chain is

$$p(k,k+1) = \frac{r}{1+r} \qquad p(k,k-1) = \frac{1}{1+r}$$

If we let $\varphi(x) = r^{-x}$ then it is well-known that if $a < x < b$ then

$$P_x(T_a < T_b) = \frac{\varphi(b) - \varphi(x)}{\varphi(b) - \varphi(a)} \qquad P_x(T_b < T_a) = \frac{\varphi(x) - \varphi(a)}{\varphi(b) - \varphi(a)} \tag{3.3.1}$$

see e.g., page 271 in Durrett (2005). From this it follows that if $\xi_0 = A$ then

$$P_A(T_0 < \infty) = r^{-|A|}.$$

Let $a = 0$, $x = |A|$, and $b \to \infty$ in the first formula in (3.3.1).

If we condition on the voter model dying out and let $h(m) = r^{-m}$ then the embedded discrete time chain in the conditioned process has

$$\bar{p}(k,k+1) = \frac{p(k,k+1)h(k+1)}{h(k)} = \frac{1}{1+r} \qquad \bar{p}(k,k-1) = \frac{r}{1+r} \tag{3.3.2}$$

In words, conditioning a super-critical biased voter model to die out turns it into a subcritical biased voter model with the parameters reversed. Thus in $(\xi_t^0 | T_0 < \infty)$ if $\xi_t^0 = A$ with $|A| = k$ and $|\partial A| = \ell$, then jumps to size $k+1$ occur at rate ℓ and to $k-1$ occur at rate $r\ell$.

Using the reasoning for (3.2.1)

$$E\left(\int_0^{T_0} |\xi_t^0| \, dt \,\middle|\, T_0 < \infty \right) = \sum_{k=1}^{\infty} \frac{\bar{P}_1(T_k < T_0)}{\bar{P}_k(T_k^+ = \infty)} \cdot \frac{k}{q(k)} \tag{3.3.3}$$

where the bar indicates conditioning on $T_0 < \infty$. By symmetry and (3.3.1)

$$\bar{P}_1(T_k < T_0) = P_{k-1}(T_0 < T_k) = \frac{r^{-k} - r^{-(k-1)}}{r^{-k} - 1} = \frac{r-1}{r^k - 1}$$

Using symmetry and (3.3.1) again,

$$\begin{aligned}\bar{P}_k(T_k^+ = \infty) &= \frac{r}{1+r}\bar{P}_{k-1}(T_0 < T_k) \\ &= \frac{r}{1+r}P_1(T_k < T_0) = \frac{r}{1+r}\cdot\frac{r^{-1}-1}{r^{-k}-1},\end{aligned}$$

so we have

$$\frac{\bar{P}_1(T_k < T_0)}{\bar{P}_k(T_k^+ = \infty)} = \left(\frac{r-1}{r^k-1}\right)\left(\frac{1+r}{r-1}\right)(1-r^{-k}) = r^{-k}(1+r)$$

and (3.3.3) becomes

$$E\left(\int_0^{T_0} |\xi_t^0|\,dt \,\middle|\, T_0 < \infty\right) = (1+r)\sum_{k=1}^{\infty} r^{-k}\cdot\frac{k}{q(k)} \tag{3.3.4}$$

One Dimension. In one dimension $q(k) = 2$, so the mean in (3.3.4) is

$$\frac{1+r}{2}\sum_{k=1}^{\infty} kr^{-k} \le \frac{1+r}{2}\left(1-\frac{1}{r}\right)^{-2} \le C(r-1)^{-2}$$

where the first inequality comes from thinking about the mean of the geometric distribution, and we have simplified in the second because we only care about how the constant blows up as $r \downarrow 1$. Since we expect $O((r-1)^{-1})$ unsuccessful attempts before finding the first family of type 1's that lives forever, (A3) is satisfied when $(r-1)^{-3} \ll 1/u_2$ or $r - 1 \gg u_2^{1/3}$.

3.3.1 Heuristics in $d \ge 2$

If one ignores some annoying details, the proof above extends easily to $d \ge 2$. By Cox, Durrett, and Perkins (2000) in $d \ge 3$, $q(k) \sim 2d\beta_d k$ for large k, so the mean

in (3.3.4) is

$$\frac{1+r}{2d\beta_d}\sum_{k=1}^{\infty} r^{-k} \le C(r-1)^{-1} \tag{3.3.5}$$

Taking into account the expected number of unsuccessful attempts, (A3) is satisfied when $(r-1)^{-2} \ll 1/u_2$ or $r-1 \gg u_2^{1/2}$.

In $d = 2$ we have $q(k) \sim 4\beta_2 k/\log k$ so the mean is

$$\frac{1+r}{8\beta_2}\sum_{k=1}^{\infty} r^{-k}\log k \le C(r-1)^{-1}\log(1/(r-1)) \tag{3.3.6}$$

and (A3) is satisfied if when $(r-1)^{-2}\log(1/(r-1)) \ll 1/u_2$ or $r-1 \gg u_2^{1/2}\log^{1/2}(1/u_2)$.

3.3.2 Rigorous proofs in $d \ge 2$

We need to estimate the expected number of man-hours in a subcritical biased voter model. We want to use interacting particle system duality, so we have to interchange the roles of 0's and 1's to get a supercritical process. Then at time $t = 0$, we place a type-1 at every lattice point except the origin, where we have a single 0. Let ζ_t^0 be the set of sites occupied by 1's in this supercritical process. The supercritical biased voter model is dual to a coalescing branching random walk $\hat{\zeta}_t$ in which particles jump to a randomly chosen neighbor at rate 1, give birth onto a randomly chosen neighbor at rate $r-1$, and two particles that land on the same site coalesce to 1. Given this framework, we reason about the evolution of the lone zero at the origin. Duality tells us that

$$P(x \notin \zeta_t^0) = P(\hat{\zeta}_t^x \subseteq \{0\})$$

If $Z_t(x) = 1-\zeta_t^0(x)$, then $E\left[\sum_x Z_t(x)\right]$ is the expected size of the patch of type-0's

at time t. By translation invariance,

$$E\left[\sum_x Z_t(x)\right] = \sum_x P(\hat{\zeta}_t^x \subseteq \{0\}) = \sum_x P(\hat{\zeta}_t^0 \subseteq \{-x\}) = P(|\hat{\zeta}_t^0| = 1) \tag{3.3.7}$$

In $d \geq 3$, random walks branch at rate $r-1$ and return together with probability $1-\beta_3$. If we do not permit $\hat{\zeta}_t^0$ to branch again while $|\hat{\zeta}_t^0| > 1$, then branches take an expected time of $1/(r-1)$ to occur, and the number of times $\hat{\xi}_t^0$ branches before two branches never again coalesce is geometric mean $1/\beta_d$. Thus

$$\int_0^\infty P(|\hat{\zeta}_t^0| = 1)\, dt \leq \frac{1}{(r-1)\beta_3}$$

which proves (3.3.5).

In $d = 2$ we begin with a crude argument which gives an upper bound of the wrong order of magnitude. Since $q(k) \geq ck^{1/2}$, using (3.3.4) gives

$$\leq (1+r)\sum_{k=1}^{\infty} r^{-k}k^{1/2} \leq C(r-1)^{-1/2}$$

To get an upper bound with the correct behavior as $r \to 1$ we use a result from Durrett and Zähle (2007). Let $\beta = r-1$ and $h(\beta) = (1/\beta)\log(1/\beta)$. We use this notation to make it easier to connect with the result in the paper cited. The reader should not confuse this with $h_d(u_2)$ defined in (3.1.4). Lemma 2.1 on page 1756 shows that if we run time at rate $h(\beta)$ and scale space by $h(\beta)^{1/2}$ then the dual process $\hat{\zeta}_t^0$ converges to a branching Brownian motion Y_t in which new particles are born at rate $\gamma = \pi$, see (b) on page 1758. In the limit process there is only branching and no coalescence, so $P(|Y_t| = 1) = e^{-\gamma t}$. Using the limit theorem now

$$P(|\hat{\zeta}_{th(\beta)}^0| = 1) \to e^{-\gamma t}$$

and it follows that

$$\frac{1}{h(r-1)}\int_0^\infty P(|\hat{\zeta}_t^0| = 1)\, dt \to 1/\gamma$$

which shows that (3.3.6) holds for r near 1.

3.4 References

Armitage, P., and Doll, R. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Brit. J. Cancer.* 8, 1–12

Armitage, P., and Doll, R. (1957) A two-stage theory of carcinogenesis. *Brit. J. Cancer.* 11, 161–169

Bramson, M., Cox T., Le Gall J. (2001) Super-Brownian Limits of Voter Model Clusters. *Ann. Probab.* Vol. 29 (3), pp. 1001-1032

Bramson, M., and Griffeath, D. (1980a) *Asymptotics for interacting particle systems on ZZ^d.* Z. Wahrscheinlichkeitstheorie verw. Gebiete. 53, 183–196

Bramson, M., and Griffeath, D. (1980b) On the Williams-Bjerknes tumour growth model. II. *Math. Proc. Cambridge Philos. Soc.* 88, 339–357

Bramson, M., and Griffeath, D. (1981) On the Williams-Bjerknes tumour growth model. I. *Ann. Probab. 9*, 173–185

Clifford, P., and Sudbury, A. (1973) A model for spatial conflict. *Biometrika.* 60, 581–588

Cox, J.T., Durrett, R., and Perkins, E.A. (2000) Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* 28, 185–234

Durrett, R. (1996) *Stochastic Calculus: A Practical Introduction*. CRC Press

Durrett, R., and Schmidt, D. (2008) Waiting for two mutations: with applica-

tions to regulatory sequence evolution and the limits of Darwinian evolution. *Genetics.* 180 (2008), 1501–1509

Durrett, R., Schmidt, D., and Schweinsberg, J. (2009). A waiting time problem arising from the study of multi-stage carcinogenesis. *Ann. Appl. Probab.* 19, 676–718

Durrett, R., and Zähle, I. (2007) On the width of hybrid zones. *Stoch. Proc. Appl.* 117, 1751–1763

Feller, W. (1971) *An Introduction To Probability Theory And Its Applications, Vol. II*. Third Edition, Wiley Press

Fisher, J.C. (1959) Multiple mutation theory of carcinogenesis. *Nature.* 181, 651–652

Frank, S.A. (2007) *Dynamics of Cancer: Incidence, Inheritance and Evolution.* Princeton Series in Evolutionary Biology

Holley, R. A., and Liggett, T. M. (1975) Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* 3, 643–663.

Iwasa, Y., F. Michor, and M. A. Nowak. (2004) Stochastic tunnels in evolutionary dynamics. *Genetics.* 166, 1571–1579

Iwasa, Y., F. Michor, N. L. Komarova, and M. A. Nowak. (2005) Population genetics of tumor suppressor genes. J. Theor. Biol. 233, 15–23

Knudson, A.G. (1971) Mutation and cancer: statistical study of retinoblastoma.

Proc. Natl. Acad. Sci. USA 68, 820–823

Knudson, A.G. (2001) Two genetic hits (more or less) to cancer. *Nature Reviews Cancer.* 1, 157–162

Komarova, N.L. (2006) Spatial stochastic models of cancer initiation and progression. *Bull. Math. Biol.* 68 (7), pp. 1573-1599.

Komarova, N.L., Sengupta, A., and Nowak, M.A. (2003) Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *J. Theor. Biol.* 223, 433–450

Liggett, T.M. (1999) *Stochastic interacting systems: contact, voter and exclusion processes.* Springer, New York.

Nordling, C.O. (1953) A new theory on cancer inducing mechanisms. *British J. Cancer.* 7, 68–72

Nowak, M. A., 2006 *Evolutionary Dynamics: Exploring the Equations of Life.* Belknap Press, Cambridge, MA

Sawyer, S. (1979) A limit theorem for patch sizes in a selectively-neutral migration model. *J. Appl. Probab.* 16, 482–495

Schweinsberg, J. (2008) The waiting time for m mutations. *Electron. J. Probab.* 13, 1442–1478

Williams, T., and Bjerknes, R. (1972) Stochastic model for abnormal clone spread through epithelial basal layer. *Nature.* 235, 19–21

Wodarz, D., and Komarova, N.L. (2005) *Computational Biology of Cancer: Lecture notes and mathematical modeling*. World Scientific